

BERICHTIGTE FASSUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
15. November 2001 (15.11.2001)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 01/85989 A2

(51) Internationale Patentklassifikation<sup>7</sup>: C12Q 1/68

(21) Internationales Aktenzeichen: PCT/EP01/05363

(22) Internationales Anmeldedatum:  
10. Mai 2001 (10.05.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
100 23 130.6 11. Mai 2000 (11.05.2000) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. [DE/DE]; Leonrodstrasse 54, 80636 München (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): RUPP, Steffen [DE/DE]; Oberer Bauernwaldweg 52, 70195 Stuttgart (DE). JOHANNES, Franz-Josef [DE/DE]; Untere Burghalde 23, 71229 Leonberg (DE). SOHN, Kai [DE/DE]; Akazienweg 6, 73527 Schwäbisch-Gmünd (DE).

(74) Anwälte: SCHRELL, Andreas usw.; Maybachstrasse 6A, 70469 Stuttgart (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht: — ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

(48) Datum der Veröffentlichung dieser berichtigten Fassung: 18. April 2002

(15) Informationen zur Berichtigung: siehe PCT Gazette Nr. 16/2002 vom 18. April 2002, Section II

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: HYPHA-SPECIFIC FACTORS FROM CANDIDA ALBICANS

(54) Bezeichnung: HYPHENSPEZIFISCHE FAKTOREN AUS CANDIDA ALBICANS

(57) Abstract: The invention relates to biochips, especially nucleotide chips that contain nucleotide sequences encoding hypha-specific proteins, to protein chips containing hypha-specific proteins, and to antibody chips that contain antibodies directed against said hypha-specific proteins. The invention further relates to diagnostic compositions that contain said nucleotide, protein or antibody chips, to methods for detecting and identifying substances that are therapeutically effective against candida species caused diseases, and to a method of diagnosing a candida caused disease.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft Biochips, insbesondere Nucleotid-Chips, die hyphenspezifische Proteine codierende Nucleotidsequenzen enthalten, Protein-Chips, die hyphenspezifische Proteine enthalten, und Antikörper-Chips, die gegen diese hyphenspezifischen Proteine gerichtete Antikörper enthalten, diagnostische Zusammensetzungen, die diese Nucleotid-, Protein- oder Antikörper-Chips enthalten, Verfahren zum Auffinden und Identifizieren von Substanzen, die gegen durch Candida-Arten verursachte Krankheiten therapeutisch wirksam sind, und Verfahren zum Diagnostizieren einer durch Candida verursachten Krankheit.

WO 01/85989 A2

## Hyphenspezifische Faktoren aus Candida albicans

### Beschreibung

Die vorliegende Erfindung betrifft Biochips, insbesondere Nucleotid-Chips, die hyphenspezifische Proteine codierende Nucleotidsequenzen enthalten, Protein-Chips, die hyphenspezifische Proteine enthalten, und Antikörper-Chips, die gegen diese hyphenspezifischen Proteine gerichtete Antikörper enthalten, diagnostische Zusammensetzungen, die diese Nucleotid-, Protein- oder Antikörper-Chips enthalten, Verfahren zum Auffinden und Identifizieren von Substanzen, die gegen durch Candida-Arten verursachten Krankheiten therapeutisch wirksam sind, und Verfahren zum Diagnostizieren einer durch Candida verursachten Krankheit.

Zu den Sprosspilzen oder Hefen zählen neben den schon seit langem zum Beispiel in der Lebensmittelindustrie kommerziell genutzten Hefen der Familie Saccharomycetaceae, auch asporogene Hefen wie beispielsweise Hefen der Gattung Candida. Einige Angehörige der Gattung Candida sind in der Lage, Mycelverbände zu bilden, andere vermehren sich lediglich durch Sprossung. Candida albicans ist der am häufigsten isolierte humanpathogene Pilz. Candida albicans verursacht häufig opportunistische Infektionen, also Infektionen durch normalerweise relativ unproblematische Keime bei immunsupprimierten Patienten. Derartige Infektionen nehmen bei diesen Patienten einen schweren Verlauf und verkürzen die

Überlebenszeit zum Beispiel HIV-Infizierter oder mittels Chemo- oder Radiotherapie behandelter Krebspatienten entscheidend. Gegenwärtig wird die Behandlung von systemischen Infektionen mit *Candida albicans* hauptsächlich mittels Azolen oder Polyenen durchgeführt. Die Behandlung mittels dieser beiden Substanzklassen weist jedoch Nachteile auf. Polyene führen zu starken Nebenwirkungen, gegen die Azole entwickeln sich zunehmend Resistenzen (DiDomenico, 1999, Curr Opin Microbiol 2, 509 bis 515, Georgopadakou, 1998, Curr Opin Microbiol 1, 547 bis 557).

Da die klinischen Befunde bei Pilzinfektionen überwiegend uncharakteristisch sind, gestaltet sich die exakte Diagnose pilzlicher Infektionen, insbesondere von *Candida*-Infektionen, äußerst schwierig. Bei Verdacht auf *Candida*-Befall der Haut oder der Schleimhäute müssen beispielsweise Oberflächenabstriche abgenommen und mikroskopisch untersucht werden. Bei Befall innerer Organe ist es erforderlich, Organbiopsien histologisch zu untersuchen, um invasives Wachstum nachweisen zu können. Zur Diagnostik einer generalisierten *Candida*-Infektion ist die Spiegelung des Augenhintergrundes indiziert. Ferner müssen mehrere Blutkulturen untersucht werden, die an aufeinanderfolgenden Tagen venös abzunehmen sind. Bei einer Nierenbeteiligung ist zusätzlich der Urin zu untersuchen. Derartige mikroskopische Nativpräparate erlauben jedoch nur den Nachweis von polymorphen Pilzzellen (Hyphen, Pseudohyphen und Blastosporen) und Sporen, ohne dass jedoch die genaue Spezies ermittelt werden kann und geeignete therapeutische Maßnahmen eingeleitet werden können.

Neben dem mikroskopischen Nachweis ist es daher unerlässlich, Kulturen zur exakten Speziesbestimmung anzulegen. Eine weitere diagnostische Möglichkeit, die jedoch bislang nicht den erhofften Aussagewert hat, ist der Nachweis von Candida-Antigen im Serum des Patienten. Ein hoher Titer spricht zwar für eine systemische Candida-Infektion, ist aber nicht beweisend, während ein negativer Befund eine systemische Infektion nicht ausschließen kann.

Die Entwicklung weiterer verbesserter Diagnostika zur zweifelfreien Zuordnung einer gesundheitlichen Störung zu den durch Vertreter der Familie Candida hervorgerufene Infektionen und von Antimycotica zur Behandlung von durch Vertreter der Familie Candida hervorgerufene Infektionen ist daher dringend erforderlich.

Das der vorliegenden Erfindung zugrundeliegende technische Problem besteht also darin, Mittel und Verfahren zur Diagnose von durch Candida albicans hervorgerufenen Infektionen und zur Entwicklung von Substanzen, die gegen Candida-verursachte Erkrankungen therapeutisch wirksam sind, zur Verfügung zu stellen.

Die Erfindung löst das ihr zugrundeliegende technische Problem durch die Bereitstellung von Biochips, insbesondere eines Nucleotid-Chips, umfassend einen festen Träger und mindestens eine daran fixierte Nucleotidsequenz, welche für die Identifizierung und Transkription eines Gens codierend für ein hyphenspezifisches Protein aus Candida, insbesondere Candida albicans, geeignet ist, wobei diese Nuc-

leotidsequenz ausgewählt ist aus der Gruppe bestehend aus:

- (a) einer Nucleotidsequenz, definiert in SEQ ID Nr. 1, 2, 3, 4, 12, 13, 15 oder 17, oder eines komplementären Strangs oder Teils davon,
- (b) einer Nucleotidsequenz, codierend eine Aminosäuresequenz, definiert in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18, oder eines komplementären Strangs oder Teils davon und
- (c) einer Nucleotidsequenz, die mit einer der (a) oder (b) genannten Nucleotidsequenzen hybridisiert.

Im Zusammenhang mit der vorliegenden Erfindung wird unter einem Biochip eine Vorrichtung verstanden, die eine Vielzahl biologischer Substanzen, beispielsweise Nucleotidsequenzen, Proteine oder Antikörper, in immobilisierter oder fixierter Form umfasst und mit deren Hilfe mittels Hybridisierungs- und/oder Bindungsverfahren eine kleine Menge eines Liganden, der unter geeigneten Bedingungen an die biologische Substanz binden kann, in einer kleinen Probe nachgewiesen werden kann. Unter einem Nucleotid-Chip wird eine Vorrichtung verstanden, die eine Vielzahl verschiedener Nucleinsäuren oder Nucleotidsequenzen wie DNA oder RNA in immobilisierter Form enthält und mit deren Hilfe mittels Nucleinsäure-Hybridisierung eine kleine Menge einer komplementären Nucleinsäure in einer kleinen Proben-

flüssigkeit oder mittels DNA/Protein-Bindungsuntersuchungen eine kleine Menge eines an Nucleinsäuren bindenden Proteins nachgewiesen werden kann.

Die erfindungsgemäßen Nucleotid-Chips enthalten an einem festen Träger fixierte Nucleotidsequenzen, die ausschließlich in der hyphal wachsenden Form von *Candida albicans* exprimierte Proteine codieren beziehungsweise die ausschließlich die Expression hyphenspezifischer Proteine regulieren. Das heisst, die auf den erfindungsgemäßen Nucleotid-Chips enthaltenen Nucleotidsequenzen werden während des hefeartigen Wachstums von *Candida* nicht exprimiert. Die erfindungsgemäß beschriebenen Nucleotidsequenzen und die davon codierten Proteine weisen keine signifikanten Homologien, beispielsweise mit Proteinen von *Saccharomyces cerevisiae*, einem verwandten, nicht-pathogenen, nicht hyphal wachsenden Pilz auf. Es ist bekannt, dass das filamentöse Wachstum, also die Bildung von Hyphen, eine wichtige Voraussetzung für die Ausprägung der Virulenzeigenschaften von *Candida* ist (Mitchell, 1998, Curr. Opin. Microbiol., 1, 687-692). So sind *Candida albicans*-Formen, die keine Hyphen ausbilden, im Modellsystem (*Mus musculus*) avirulent (Lo et al., 1997, Cell 90, 939 bis 949). Im Zusammenhang mit der vorliegenden Erfindung wird daher unter einem hyphenspezifischen Protein ein Protein und/oder Peptid verstanden, das ausschließlich in Arten der Gattung *Candida* exprimiert wird und vorzugsweise für die Virulenz von *Candida*, insbesondere *Candida albicans* Bedeutung hat.

Die erfindungsgemäß verwendeten, spezifisch in der pathogenen Form von *Candida albicans* vorkommenden Nucleotidsequenzen und die von diesen codierten Proteine stellen daher ausgezeichnete diagnostische Hilfsmittel für das Erkennen lokaler oder systemischer Candidosen dar, insbesondere zum Erkennen lokaler oder systemischer *Candida albicans*-Infektionen. Darüber hinaus bieten sie die Möglichkeit, im Falle einer *Candida*-Infektion hyphal wachsende, also virulente *Candida albicans*-Formen von hefeartig wachsenden, also nicht-virulenten *Candida albicans*-Formen zu unterscheiden.

Überdies erweisen sich die erfindungsgemäß verwendeten Nucleotidsequenzen und Proteine als besonders wertvoll für die Entwicklung von Medikamenten zur Bekämpfung von Candidosen. Die erfindungsgemäßen Nucleotidsequenzen und Proteine können als Targets für die Identifikation spezifisch auf diese wirkender Substanzen eingesetzt werden. So können etwa Substanzbibliotheken auf die Interaktion der in ihnen vorhandenen Substanzen mit erfindungsgemäßen Proteinen oder erfindungsgemäßen Nucleotidsequenzen hin untersucht werden.

Die Erfindung betrifft daher in bevorzugter Ausführungsform einen vorgenannten Nucleotid-Chip, der eine Nucleotidsequenz umfasst, die eine proteincodierende Nucleotidsequenz ist, ausgewählt aus der Gruppe bestehend aus den Nucleotidsequenzen der SEQ ID Nr. 1, 2, 3, 4, 13, 15 und 17. Diese Nucleotidsequenzen codieren die in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäuresequenzen. Diese Sequenzen sind besonders hilfreich bei der

Diagnose von Erkrankungen, die von Candida-Arten verursacht werden, indem sie den gezielten Nachweis der Gegenwart von Candida in Oberflächenabstrichen oder Organbiopsien ermöglichen. Beispielsweise können die auf einem solchen Nucleotid-Chip enthaltenen erfindungsgemäßen Nucleotidsequenzen mit markierten DNA-Proben, die aus Quellen wie Hautabstrichen, Biosieproben oder eigens angelegten Pilzkulturen isoliert oder mittels PCR-Verfahren amplifiziert wurden, unter stringenten Bedingungen hybridisiert werden. Da die erfindungsgemäßen Nucleotidsequenzen keine signifikanten Homologien mit Nucleotidsequenzen von verwandten Pilzen aufweisen, erlaubt der Nachweis einer Hybridisierung daher den Nachweis von Candida in einer mit Pilzen infizierten Probe. Die erfindungsgemäß verwendeten Nucleotidsequenzen ermöglichen jedoch nicht nur den einfachen Nachweis von Candida, sondern auch den Nachweis, dass Candida in hyphaler Form wächst und dementsprechend virulente Eigenschaften aufweist. Beispielsweise kann aus Hautabstrichen oder Biosieproben gezielt mRNA isoliert und/oder mittels PCR-Verfahren amplifiziert werden. Nach Markierung mit geeigneten Markierungsmitteln wird die so erhaltene mRNA mit dem die erfindungsgemäßen Nucleotidsequenzen enthaltenden Nucleotid-Chip hybridisiert. Da die erfindungsgemäßen Nucleotidsequenzen ausschließlich während des hyphalen Wachstums von Candida transkribiert und exprimiert werden, nicht jedoch während des hefeartigen Wachstums, erlaubt der Nachweis einer Hybridisierung unter Verwendung von isolierter mRNA den Nachweis, dass sich Candida den Übergang in der hyphalen und damit virulenten Wachstumsphase befindet.



Die Erfindung betrifft in einer weiteren bevorzugten Ausführungsform den vorgenannten Nucleotid-Chip, der Nucleotidsequenzen enthält, die regulatorische Elemente von hyphenspezifische Candida-Proteine codierenden Genen darstellen, also Elemente, die insbesondere die Transkription der mit diesen regulatorischen Elementen funktionell verbundenen proteincodierenden Bereiche ermöglichen, beispielsweise Promotoren, Transkriptionsterminationssignale, Silencer, Enhancer usw.. Diese besonders bevorzugten Nucleotidsequenzen können insbesondere Promotoren sein, besonders bevorzugt der in SEQ ID Nr. 12 dargestellte Promotor. Die erfindungsgemäßen regulatorischen Elemente, insbesondere Promotoren, besonders bevorzugt der in SEQ ID Nr. 12 dargestellte Promoter, erweisen sich insofern als besonders vorteilhaft, als sie die für die Induktion von hyphenspezifisch exprimierten Proteinen notwendigen Regulationssequenzen umfassen und dementsprechend verwendet werden können, um weitere spezifische Candida-Proteine zu identifizieren, die in vivo durch Bindung an diese regulatorischen Elemente die Transkription von hyphenspezifischen Proteinen, insbesondere der Proteine mit den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäuresequenzen, induzieren oder hemmen können. Nach Identifizierung solcher Proteine können Nucleotid-Chips, die regulatorische Elemente von hyphenspezifisch exprimierten proteincodierenden Nucleotidsequenzen enthalten, verwendet werden, um Substanzen jedweder Art zu identifizieren, die die Wechselwirkung zwischen den regulatorischen Elementen und den daran bindenden Proteinen inhibieren. Unter Verwendung derartiger Nucleotid-Chips ist es also mög-

lich, Substanzen zu identifizieren, die die Expression hyphenspezifischer Proteine hemmen und daher potentiell als spezifisch wirkende Medikamente gegen Candida-Infektionen eingesetzt werden können.

Die Erfindung betrifft in einer weiteren, besonders bevorzugten Ausführungsform den vorgenannten Nucleotid-Chip, der als Nucleotidsequenz DNA-, RNA- oder PNA-Sequenzen aufweist. Bei PNA (Peptide Nucleic Acid oder Polyamide Nucleic Acid)-Sequenzen handelt es sich um Moleküle, die nicht negativ geladen sind und in gleicher Weise wie DNA wirken (Nielsen et al., 1991, Science, 254, 1497-1500; Nielsen et al., 1997, Biochemistry, 36, 5072-5077; Weiler et al., 1997, Nuc. Acids Res., 25, 2792-2799). PNA-Sequenzen umfassen ein Polyamid-Grundgerüst aus N-(2-Aminoethyl)-glycin-Einheiten und besitzen keine Glucose-Einheiten und keine Phosphat-Gruppen.

Die erfindungsgemäßen Nucleinsäuremoleküle, die an einem Träger fixiert werden können, können aus natürlichen Quellen, vorzugsweise aus Candida albicans, isoliert werden. Beispielsweise können die Nucleinsäuremoleküle mittels PCR-Verfahren isoliert und amplifiziert werden, wobei doppelsträngige Moleküle erhalten werden. Die erfindungsgemäßen Nucleinsäuremoleküle können aber auch nach bekannten Verfahren in vitro synthetisiert werden, wobei einzelsträngige Oligonucleotide oder Peptid-Oligonucleotide erhalten werden. Durch die Wahl geeigneter Primer können gewünschte Bereiche der erfindungsgemäßen Nucleinsäuren, das heißt sowohl einzelne Bereiche als auch der gesamte Leseraster des Gens, amplifiziert und isoliert werden. Mittels

gängiger molekularbiologischer Techniken ist es möglich, verschiedenartige Mutationen in die erfindungsgemäßen Nucleinsäuremoleküle einzufügen. Dadurch können beispielsweise Sequenzvarianten erfasst werden, die in unterschiedlichen klinischen Candida-Isolaten vorkommen. Derartige von der Erfindung erfasste Mutationen können Insertionen, Deletionen, Duplikationen, Inversionen, Additionen, Austausche oder ähnliches sein, auch von ungewöhnlichen Nucleotiden. Auf diese Weise können aber auch modifizierte Oligonucleotide mit funktionellen Gruppen hergestellt werden, die eine kovalente Bindung des Oligonucleotids an das Trägermaterial zur Herstellung des erfindungsgemäßen Nucleotid-Chips ermöglichen. So können beispielsweise Oligonucleotide mit Amino-Modifikationen oder Biotin-Gruppen hergestellt werden, die an auf der Oberfläche des Trägermaterials enthaltenen chemisch reaktiven Gruppen (Epoxide) oder Streptavidin-Gruppen oder Derivate davon kovalent binden können. In einer weiteren bevorzugten Ausführungsform der Erfindung ist vorgesehen, dass Nucleinsäuren mit photolabile Schutzgruppen enthaltenden Nucleosid-Derivaten versehen werden.

Erfindungsgemäß können für den Nucleotid-Chip auch Nucleotidsequenzen verwendet werden, die durch Fusion der erfindungsgemäßen Nucleotidsequenzen mit Genen oder Bestandteilen von Genen aus anderen Quellen erzeugt werden. Erfindungsgemäß können auch verkürzte Nucleotidsequenzen der vorgenannten Art verwendet werden, sofern diese die genannte Hyphenspezifität aufweisen. Erfindungsgemäß ist

vorgesehen, dass verkürzte Nucleotidsequenzen eine Länge von mindestens 15 Basenpaaren aufweisen.

In einer weiteren Ausführungsform der Erfindung ist vorgesehen, dass der Nucleotid-Chip auch Nucleotidsequenzen der vorgenannten Art umfasst, die mit den vorgenannten erfindungsgemäßen Nucleotidsequenzen hybridisieren. Hybridisierung bedeutet im Zusammenhang mit diesem Aspekt der Erfindung eine Hybridisierung unter konventionellen Hybridisierungsbedingungen, wie sie in Sambrook et al. (Molecular cloning. A laboratory manual. Cold spring harbor laboratory press, 2. Ausgabe 1989) beschrieben sind, vorzugsweise unter stringenten Bedingungen. Gemäß vorliegender Erfindung spricht man von einer Hybridisierung, wenn nach dem Waschen für eine Stunde mit 1 x SSC und 0,1% SDS bei 55°C, vorzugsweise bei 62°C und besonders bevorzugt bei 68°C insbesondere für eine Stunde 0,2 x SSC und 0,1% SDS bei 55°C, vorzugsweise bei 62°C und besonders bevorzugt bei 68°C noch ein positives Hybridisierungssignal beobachtet wird. Erfindungsgemäß kann eine unter derartigen Waschbedingungen mit einer der in den Sequenzprotokollen angegebenen Nucleotidsequenzen hybridisierende Nucleotidsequenz durch Immobilisierung an den festen Träger für den erfindungsgemäßen Nucleotid-Chip verwendet werden.

Die Identifizierung und Isolierung hybridisierender Nucleotidsequenzen kann beispielsweise unter Verwendung eines erfindungsgemäßen Nucleotid-Chips, der die vorstehend genannten Nucleotidsequenzen oder Teile dieser Moleküle beziehungsweise des komplementären Stranges enthält, erfolgen. Der zur I-

identifizierung und Isolierung hybridisierender Nucleotidsequenzen eingesetzte Nucleotid-Chip kann beispielsweise Nucleotidsequenzen enthalten, die exakt die oder im wesentlichen die unter SEQ ID Nr. 1 bis 4, 12, 13, 15 oder 17 dargestellte Nucleotidsequenzen oder Teile dieser Sequenzen oder komplementäre Stränge aufweisen. Der Nucleotid-Chip kann aber auch synthetische Fragmente enthalten, die mit Hilfe üblicher Synthesetechniken hergestellt werden und deren Sequenz im wesentlichen mit der einer erfindungsgemäßen Nucleotidsequenz übereinstimmt. Auf diese Weise können Nucleotidsequenzen aus klinischen Candida-Isolaten isoliert und für den erfindungsgemäßen Nucleotid-Chip verfügbar gemacht werden, die gegenüber den in SEQ ID Nr. 1 bis 4, 9 bis 13, 15 oder 17 dargestellten Nucleotidsequenzen Abweichungen oder Mutationen enthalten.

Die mit den erfindungsgemäßen Nucleotidsequenzen hybridisierenden Moleküle umfassen auch Fragmente, Derivate, funktionelle Äquivalente und/oder allelische Varianten der vorstehend beschriebenen Nucleotidsequenzen, die ein erfindungsgemäßes Protein codieren oder dessen hyphenspezifische Expression gewährleisten. Unter „Fragmenten“ werden dabei Teile der Nucleotidsequenzen verstanden, die lang genug sind, um das hyphenspezifisch exprimierte Protein zu codieren oder die Hyphenspezifität zu gewährleisten. Der Ausdruck „Derivat“, „funktionelles Äquivalent“ oder „mutante Abwandlung“ bedeutet im Zusammenhang mit der vorliegenden Erfindung, dass sich die Sequenzen dieser Moleküle von den Sequenzen der vorstehend beschriebenen Nucleotidsequenzen an einer oder mehreren Positionen unterscheiden,

aber einen hohen Grad an Homologie zu diesen Sequenzen auf der Nucleotidebene aufweisen. Homologie bedeutet dabei eine Sequenzidentität von mindestens 40%, insbesondere eine Identität von mindestens 60%, vorzugsweise über 80%, und besonders bevorzugt, über 90%, 95%, 97% oder 99% auf Nucleinsäureebene.

Die erfindungsgemäßen Nucleotid-Chips umfassen an einem festen Träger fixierte oder immobilisierte Nucleotidsequenzen. Im Zusammenhang mit der vorliegenden Erfindung bedeutet der Begriff „fester Träger“ eine unlösliche Matrix. In bevorzugter Ausführungsform besteht der feste Träger aus einem hydrophoben oder schwach hydrophilen Material, wie transparentem Glas, Siliciumdioxid, Metalloxiden, Polymeren und Copolymeren von Dextranen oder Amiden, beispielsweise Acrylamid-Derivaten, Cellulose, Nylon, oder polymeren Materialien, wie Polyethylenterephthalat, Celluloseacetat, Polystyrol oder Polymethylmethacrylat oder einem Polycarbonat von Bisphenol A. Das Trägermaterial wird vorzugsweise vor Fixierung der Nucleotidsequenzen mit einem Oberflächen-aktivierenden Mittel wie Poly-L-Lysin, Polyethylenimin oder Polyalkylamin vorbehandelt, um die Fixierung der Nucleotidsequenzen am Trägermaterial zu verbessern. In einer anderen Ausführungsform wird als Träger verwendetes Glas mit einem Silan-Kupplungsmittel, das eine Amino-Gruppe, eine Aldehyd-Gruppe oder eine Epoxy-Gruppe aufweist, vorbehandelt. Für den erfindungsgemäßen Nucleotid-Chip können jedoch auch die im Handel erhältlichen, bereits beschichteten Trägertypen wie Poly-L-Lysin (Sigma Diagnostics), Super-Aldehyde (Telechem), Su-

per-Amine (Telechem), Silane Prep (Sigma), CMT GAPS (Corning), Type I (Clontech), Type II (Clontech), Arraylink (GeneScan Europe), Type I (Eppendorf), Type II (Eppendorf), Epoxysilan (Quantifoil) und Cast/FastSlides (Schleicher & Schüll) verwendet werden. Weitere geeignete Träger sind solche, die für photolithographisch hergestellte Nucleotid-Chips verwendet werden, beispielsweise die in Lipshutz et al. beschriebenen (Lipshutz, Fodor, Gingeras und Lockhart, 1999, Nat. Genet., 21, 20-24). In besonders bevorzugter Ausführungsform werden für den erfindungsgemäßen Nucleotid-Chip Träger mit Beschichtungen aus Poly-L-Lysin, wie in DeRisi et al. beschrieben (DeRisi, J. L., Iyer, V. R. und Brown, P. O., 1997, Science, 278, 680-686), beispielsweise Poly-Prep Slides (Sigma Diagnostics), oder Aminosilanen, wie Silane-Prep Slides (Sigma), CMT GAPS Slides (Corning) und Super Amine (Telechem), oder Membranen, wie CAST Slides oder FAST Slides (Schleicher & Schüll) verwendet. Zur Immobilisierung amino-modifizierter Oligonucleotide sind epoxy-modifizierte Oberflächen, wie ArrayLink Biochip (GeneScan Europe) oder Epoxysilane Slides (Quantifoil) besonders bevorzugt.

Die Nucleotidsequenzen können über chemische oder photochemische Reaktionen oder durch elektrostatische Wechselwirkungen an das Trägersubstrat gebunden und fixiert werden. In einer bevorzugten Ausführungsform der Erfindung ist vorgesehen, dass die Immobilisierung oder Fixierung der Nucleinsäuren an den verwendeten Trägersoberflächen über eine elektrostatische Bindung oder eine kovalente Bindung erfolgt. Wenn die Nucleinsäuren beispielsweise syn-

thetisch hergestellt wurden und eine funktionelle Gruppe aufweisen, können die Nucleinsäuren an geeignete funktionelle Gruppen auf der Oberfläche des Trägermaterials kovalent gebunden und fixiert werden (Lamture et al., 1994, Nucl. Acids Res., 22, 2121-2125; Guo et al., 1994, Nucl. Acids Res., 22, 5456-5465). Erfindungsgemäß können die Nucleinsäuren auch über Spacer oder ein Vernetzungsmittel, beispielsweise ein bifunktionelles Vernetzungsmittel, auf den Oberflächen-aktivierten Träger kovalent gebunden werden. In einer bevorzugten Ausführungsform der Erfindung ist vorgesehen, dass die Bindung der Nucleotidsequenzen an den Träger im Falle von Polylysin-, Aminosilan- und Membranbeschichteten Nucleotid-Chips mittels UV-Quervernetzung und im Falle von epoxy-modifizierten Chips mittels chemischer Reaktion erfolgt. Die Bindung der Nucleinsäuren an den Träger kann selbstverständlich auch über photochemische Reaktionen erfolgen. Im Falle solcher photolithographisch hergestellten Nucleotid-Chips erfolgt im Anschluss an die Immobilisierung eine gezielte Abspaltung der photolabilen Schutzgruppen mittels Photolyse.

Eine weitere bevorzugte Ausführungsform der vorliegenden Erfindung betrifft daher Verfahren zur Herstellung von Nucleotid-Chips gemäß Anspruch 1, umfassend die Isolierung und/oder Amplifikation von mindestens einer Nucleotidsequenz, die ein hyphenspezifisch exprimiertes Protein von Candida codiert und/oder Regulationselemente dieser Nucleotidsequenz umfasst, oder die chemische Synthese dieser Nucleotidsequenz, die Modifikation der Nucleotidsequenz während oder nach der Synthese oder



Amplifikation durch den Einbau funktioneller Gruppen oder Spacer-Einheiten, das Auftragen einer wässrigen Lösung der isolierten oder synthetisierten Nucleotidsequenz auf ein festes Trägermaterial und die Immobilisierung der Nucleotidsequenz an dem Träger mittels chemischer oder photochemischer Reaktion oder elektrostatischer Wechselwirkung.

Eine weitere besonders bevorzugte Ausführungsform der Erfindung betrifft Protein-Chips, umfassend einen festen Träger und mindestens ein daran fixiertes hyphenspezifisches Protein, ausgewählt aus der Gruppe bestehend aus:

- (a) einem Protein mit einer Aminosäuresequenz, definiert in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18, oder einem Fragment davon, und
- (b) einem Protein mit einer Aminosäuresequenz, die zu einer der in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 definierten Aminosäuresequenz eine Sequenzidentität von mindestens 80% aufweist, oder einem Fragment davon.

Im Zusammenhang mit der vorliegenden Erfindung wird unter einem Protein-Chip eine Vorrichtung verstanden, die eine Vielzahl verschiedener Proteine oder Peptide in immobilisierter Form enthält und mit deren Hilfe eine kleine Menge eines Liganden, beispielsweise eines Proteins oder eines Antikörpers, das/der an mindestens ein auf dem Träger fixiertes Protein oder Peptid kovalent oder nicht-kovalent

binden kann, in einer kleinen Probenflüssigkeit nachgewiesen werden kann.

Die erfindungsgemäßen Protein-Chips enthalten an einem festen Träger fixierte Proteine, die ausschließlich in der hyphal wachsenden Form von *Candida albicans* exprimiert werden, oder Teile davon. Die erfindungsgemäßen Protein-Chips können daher beispielsweise zum Nachweis von Antikörpern verwendet werden, die im Körper eines Organismus, insbesondere eines Säugers, als Folge einer Immunisierung durch Antigen-Determinanten von hyphal wachsenden *Candida*-Formen, insbesondere hyphenspezifischen *Candida*-Proteinen, gebildet wurden. Die Bindung von mindestens einem Antikörper aus Blut, Lymphe, Körpersekreten oder anderen Körperflüssigkeiten eines Organismus an den erfindungsgemäßen Protein-Chip ermöglicht also den Nachweis einer systemischen *Candida*-Infektion in diesem Organismus, die zur Bildung des gebundenen Antikörpers geführt hat. Darüber hinaus können die erfindungsgemäßen Protein-Chips auch verwendet werden, um beispielsweise solche Proteine aus *Candida*-infizierten Materialien zu identifizieren und zu isolieren, die in vivo mit den auf dem Protein-Chip enthaltenen Proteinen in Wechselwirkung treten. Nach Identifizierung und Isolierung derartiger interagierender Proteine können die erfindungsgemäßen Protein-Chips auch verwendet werden, um Substanzen jedweder Art zu identifizieren, die die Wechselwirkung zwischen den erfindungsgemäßen Proteinen und damit interagierenden Proteinen hemmen oder fördern können. Unter Verwendung der erfindungsgemäßen Protein-Chips lassen sich also Substanzen detektieren, die potentiell

- 18 -

als Medikamente zur Bekämpfung von Candida-Infektionen, insbesondere zur Hemmung des Übergangs vom hefeartigen Wachstum zum hyphalen Wachstum von Candida, geeignet sind.

In einer besonders bevorzugten Ausführungsform der Erfindung ist daher vorgesehen, dass der erfindungsgemäße Protein-Chip neben den hyphenspezifischen Proteinen mit den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäuresequenzen auch Derivate, funktionelle Äquivalente oder Varianten dieser Proteine umfassen kann. Im Zusammenhang mit der vorliegenden Erfindung werden unter „Derivaten, funktionellen Äquivalenten und Varianten“ insbesondere solche Abkömmlinge der Proteine mit den in SEQ ID Nr. 5 bis 8, 14, 16 und 18 angegebenen Aminosäuresequenzen verstanden, die unter Beibehaltung der Grundstruktur dieser Proteine durch Substitution von Atomen oder Molekülgruppen erhalten werden und deren Aminosäuresequenzen sich von den in SEQ ID Nr. 5 bis 8, 14, 16 und 18 angegebenen Aminosäuresequenzen an mindestens einer Position unterscheiden und die im wesentlichen einen hohen Grad an Homologie auf Aminosäureebene aufweisen. Der dem Fachmann bekannte Begriff "Homologie" bezeichnet den Grad der Verwandtschaft zwischen zwei Polypeptiden, der durch das Ausmaß der Übereinstimmung zwischen diesen Polypeptiden bestimmt wird. Dabei kann eine Übereinstimmung sowohl eine identische Übereinstimmung, also Sequenzidentität, als auch einen konservativen Aminosäureaustausch bedeuten. Vorzugsweise besitzen erfindungsgemäß verwendete Derivate, Varianten oder funktionelle Äquivalente eine Sequenzidentität zu jeweils einer

in SEQ ID Nr. 5 bis 8, 14, 16 und 18 angegebenen Aminosäuresequenzen von mindestens 80%, vorzugsweise 85% und besonders bevorzugt von über 90%, 95%, 97% und 99% auf Aminosäureebene. Die Abweichungen zu den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 dargestellten Aminosäuresequenzen können beispielsweise durch mit technischen Mitteln erzeugte Deletionen, Substitutionen, Insertionen, Additionen, Austausche oder Rekombinationen der die Aminosäuresequenzen codierenden Nucleotidsequenzen entstanden sein. Es kann sich dabei aber auch um natürlicherweise auftretende Variationen handeln, beispielsweise um auf natürliche Weise entstandene Aminosäuresequenz-Änderungen. Derivate oder Varianten der erfindungsgemäßen hyphenspezifischen Proteine können beispielsweise aus klinischen Isolaten von *Candida* stammen.

Solche Derivate, funktionelle Äquivalente oder Varianten können sich von den Proteinen mit den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäuresequenzen beispielsweise durch eine veränderte Stabilität, Spezifität, ein modifiziertes Temperatur-, pH-Wert- und/oder Konzentrationsprofil, eine veränderte Aktivität und/oder ein verändertes Effektorenmuster unterscheiden. Derivate, funktionelle Äquivalente oder Varianten können auch in anderen Konformationen vorkommen oder andere Untereinheiten beziehungsweise prä- und/oder posttranslationelle Modifikationen aufweisen. Trotz der möglicherweise vorhandenen Unterschiede besitzen die hyphenspezifischen Proteine mit den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäuresequenzen und Derivate, Varianten oder funk-

tionellen Äquivalente davon jedoch bestimmte gemeinsame Charakteristika, wie Aktivität, Molekulargewicht, immunologische Reaktivität, Konformation und/oder physikalische Eigenschaften, wie das Laufverhalten in Gelelektrophorese und deren Löslichkeit und andere.

In einer weiteren besonders bevorzugten Ausführungsform der Erfindung ist vorgesehen, dass der erfindungsgemäße Protein-Chip Fragmente der Proteine mit den in SEQ ID Nr. 5 bis 8, 14, 16 und 18 angegebenen Aminosäuresequenzen oder der Proteine, deren Aminosäuresequenz eine Sequenzidentität von mindestens 80% zu einer der in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 definierten Aminosäuresequenz aufweist, umfasst. Im Zusammenhang mit der vorliegenden Erfindung werden unter „Fragmenten“ insbesondere diejenigen isolierten Bereiche eines Proteins verstanden, die weniger Aminosäuren als das native Protein aufweisen, deren Länge jedoch ausreicht, dass das isolierte Fragment zumindest eine der für das native Protein charakteristischen Funktionen, wie Bindungsvermögen an ein zweites Protein, eine spezifische katalytische Aktivität etc. ausüben kann. In einer besonders bevorzugten Ausführungsform der Erfindung umfasst das Fragment eines Proteins einen Protein-Bereich, der eine Antigen-Determinante oder ein Epitop darstellt und daher in besonderem Maße zur Bindung eines Antikörpers geeignet ist.

Die erfindungsgemäßen hyphenspezifischen Proteine, insbesondere die Proteine mit den in SEQ ID Nr. 5, 6, 7, 8, 14, 16 und 18 dargestellten Aminosäurese-

quenzen, die auf dem erfindungsgemäßen Protein-Chip immobilisiert sind, können aus natürlichen Quellen, beispielsweise aus Candida-infizierten Geweben oder eigens angelegten Candida-Kulturen unter Verwendung üblicher, auf dem Fachgebiet bekannter Verfahren isoliert und aufgereinigt worden sein. Die verwendeten Proteine oder Fragmente können auch synthetischen Ursprungs sein. Beispielsweise lassen sich mit Hilfe des Verfahrens von Merrifield (1985, Angew. Chem., 97, 801) Peptide, also Fragmente der erfindungsgemäßen Proteine synthetisch herstellen. In besonders bevorzugter Ausführungsform der Erfindung können die hyphenspezifischen Proteine oder Peptide mittels üblicher DNA-Rekombinationstechniken hergestellt werden. Beispielsweise können die erfindungsgemäßen proteincodierenden Nucleotidsequenzen, wie die in SEQ ID Nr. 1, 2, 3, 4, 13, 15 und 17 dargestellten Sequenzen oder damit hybridisierende Sequenzen, unter Verwendung üblicher molekularbiologischer und gentechnischer Verfahren in geeigneten Vektoren insertiert und cloniert werden. Vorzugsweise erfolgt die Insertion der erfindungsgemäßen proteincodierenden Nucleotidsequenzen so, dass sie unter der Kontrolle regulatorischer Elemente stehen, das heißt mit diesen operativ verbunden sind. Diese regulatorischen Elemente gewährleisten die Transkription und Synthese translatierbarer Nucleinsäuremoleküle in pro- und/oder eukaryotischen Zellen. Bei regulatorischen Elementen kann es sich um Promotoren, Enhancer, Operatoren, Silencer und/oder Transkriptionsterminationssignale usw. handeln. Nach Transformation in geeignete Wirtszellen, beispielsweise prokaryotische oder eukaryotische Zellen, wie Bakterien-, Hefe-, Pflanzen-, In-

sekten- oder Säugerzellen, können diese Wirtszellen in einem geeigneten Kulturmedium unter solchen Bedingungen kultiviert werden, die die Bildung des von der proteincodierenden Nucleotidsequenz codierten hyphenspezifischen Proteins oder eines Fragments davon erlauben. Anschließend kann das Protein oder Fragment davon unter Verwendung geeigneter Verfahren aus der Wirtszelle oder dem Medium, in dem die Wirtszelle kultiviert wurde, isoliert und aufgereinigt werden. Zur Herstellung der hyphenspezifischen Proteine oder Fragmente können aber auch geeignete in vitro-Transkriptions-/Translationsysteme eingesetzt werden. In einer besonders bevorzugten Ausführungsform der Erfindung erfolgt die Herstellung der hyphenspezifischen Proteine oder Fragmente davon in bakteriellen Expressionssystemen, wobei die Proteine vorzugsweise als GST-Fusionsproteine, HIS-Tag-Fusionsproteine, pMAL-Fusionsproteine usw. erhalten werden.

Als fester Träger für die erfindungsgemäßen Protein-Chips können die gleichen Materialien verwendet werden, wie vorstehend für die erfindungsgemäßen Nucleotid-Chips beschrieben, beispielsweise Glas, Siliciumdioxid, andere Silica-Materialien, polymere Materialien, wie Fluorpolymere, oder Metalloxide. Vorzugsweise werden diese Trägermaterialien vor der Immobilisierung der Proteine vorbehandelt, beispielsweise mit Silan-Kupplungsmitteln. In einer bevorzugten Ausführungsform der Erfindung werden die von Joos et al. beschriebenen Epoxy-modifizierte Träger oder Membranen (Joos et al., 2000, Electrophoresis, 21, 2641-2650) verwendet.

Erfindungsgemäß ist vorgesehen, dass die Bindung und Immobilisierung der hyphenspezifischen Proteine oder Fragmente am Trägermaterial mittels chemischer oder photochemischer Reaktion oder elektrostatischer Wechselwirkung erfolgt. Die erfindungsgemäßen hyphenspezifischen Proteine und Fragmente davon können beispielsweise durch eine Vielzahl üblicherweise verwendeter funktioneller Gruppen und/oder Spacer oder chemischer Vernetzungsmittel, wie bifunktioneller Vernetzungsmittel, am Trägermaterial gebunden und immobilisiert werden. Eine Übersicht über geeignete funktionelle Gruppen, die eine Bindung von Proteinen an silanisierte Oberflächen ermöglichen, findet sich beispielsweise in Weetall, 1996, *Advances in Molecular and Cell Biology*, Bd. 15A, 161-192, JAI Press Inc. Falls das zu immobilisierende Protein als GST-Fusionsprotein vorliegt, kann die Bindung des Proteins an den Träger über auf der Trägeroberfläche vorhandene GSH-Einheiten erfolgen. Die Immobilisierung eines pMAL-Fusionsproteins kann über MBP-Einheiten auf der Oberfläche des Trägermaterials erfolgen. Liegt das zu immobilisierende Protein als HIS-Tag-Fusionsprotein, kann die Immobilisierung über  $\text{Ni}^{2+}$ -Nitrilotriacetic Acid-Oberflächen (Ni-NTA) erfolgen (Adachi et al., *Proc. Nat. Acad. Sci. USA*, 97, 7243-7247).

Eine weitere bevorzugte Ausführungsform der vorliegenden Erfindung betrifft daher Verfahren zur Herstellung von Protein-Chips, umfassend die Isolierung von mindestens einem hyphenspezifisch exprimierten Candida-Protein aus einer geeigneten Quelle oder die chemische Synthese oder rekombinante Her-



stellung dieses Proteins oder eines Fragments davon, die Modifikation des Proteins oder Fragments, während oder nach der Isolierung, Synthese oder Herstellung durch den Einbau funktioneller Gruppen oder Spacer-Einheiten, das Auftragen einer wässrigen Lösung des isolierten oder synthetisierten Proteins auf ein festes Trägermaterial und die Immobilisierung des Proteins an dem Träger mittels chemischer oder photochemischer Reaktion oder elektrostatischer Wechselwirkung.

Eine besonders bevorzugte Ausführungsform der vorliegenden Erfindung betrifft einen Antikörper-Chip, umfassend einen festen Träger und mindestens einen daran fixierten Antikörper, der spezifisch gegen ein Protein mit der in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 dargestellten Aminosäuresequenz oder ein Fragment davon oder ein Derivat davon gerichtet ist.

Da die auf dem erfindungsgemäßen Antikörper-Chip fixierten Antikörper spezifisch gegen hyphenspezifische Candida-Proteine gerichtet sind, kann unter Verwendung eines derartigen Chips die Gegenwart von hyphal wachsenden Candida-Zellen in pilzinfizierten Haut- oder Schleimhautabstrichen, Organbiopsien oder Körperflüssigkeiten nachgewiesen werden. Beispielsweise können aus den vorstehend genannten Proben Proteine extrahiert und nach Markierung mit dem erfindungsgemäßen Antikörper-Chip inkubiert werden. Die Bindung mindestens eines markierten Proteins an den erfindungsgemäßen Antikörper-Chip zeigt daher an, dass in der untersuchten Probe hyphal wachsende Candida-Formen vorliegen.

Im Zusammenhang mit der vorliegenden Erfindung wird unter einem „Antikörper“ ein Polypeptid verstanden, das durch ein oder mehrere Immunglobulin-Gene codiert wird und spezifische Strukturen auf einem Antigen, insbesondere eine Antigen-Determinante oder ein Epitop, erkennt und spezifisch daran binden kann. Der Begriff „Antikörper“ umfasst nicht nur ein vollständiges Immunglobulin, sondern auch eine Reihe von Fragmenten, die mittels Spaltung mit verschiedenen Peptidasen erhältlich sind. Der Begriff „Antikörper“ umfasst auch modifizierte Antikörper, wie oligomere, reduzierte, oxidierte und markierte Antikörper. „Antikörper“ umfasst auch Antikörper-Fragmente, die sowohl durch Modifikation intakter Antikörper als auch unter Verwendung von DNA-Rekombinationstechniken erzeugt wurden. Im Zusammenhang mit der vorliegenden Erfindung umfasst „Antikörper“ insbesondere auch solche Fragmente wie Fab,  $F(ab')_2$  oder Fvm, die an eine Antigen-Determinante binden können. Das Fab-Fragment kann durch Spaltung des intakten Antikörpers mit dem Enzym Papain erzeugt werden, wobei eine intakte leichte Kette mit einem Teil einer schweren Kette erhalten wird.  $F(ab')_2$  kann durch Behandlung des intakten Antikörpers mit Pepsin ohne anschließende Reduktion erzeugt werden. Bei  $F(ab')_2$  handelt es sich um ein aus zwei Fab'-Fragmenten bestehendes Dimer. Bei Fv handelt es sich um ein gentechnisch hergestelltes Antikörper-Fragment, das den variablen Bereich der leichten Kette und den variablen Bereich der schweren Kette umfasst. Verfahren zur Herstellung solcher Fragmente sind beispielsweise von Harlow und Lane in „Antibodies: A Laboratory

Manual", 1988, Cold Spring Harbor Laboratory, New York, beschrieben.

Der Ausdruck „Antikörper, der spezifisch gegen ein Protein gerichtet ist“ oder „Antikörper, der spezifisch an ein Protein bindet“ bedeutet, dass ein Antikörper unter definierten Immuntest-Bedingungen eine Antigen-Determinante oder ein Epitop eines Proteins erkennen und mittels seines Paratops daran binden kann. Antigen-Determinanten bestehen meist aus chemisch aktiven Molekül-Gruppen, wie Aminosäuren oder Zuckerseitenketten, auf der Oberfläche eines Antigens, beispielsweise eines Proteins, und besitzen eine charakteristische dreidimensionale Struktur. Unter definierten Bedingungen bindet ein Antikörper daher vorzugsweise nur an ein bestimmtes Protein, während an andere Proteine in der gleichen Probe keine nennenswerte Bindung erfolgt.

Ein erfindungsgemäß auf einem Antikörper-Chip immobilisierter Antikörper kann daher an ein Protein, Peptid, Kohlenhydrat, Proteoglycan und/oder einen Lipidkomplex binden, das/der mit den erfindungsgemäß verwendeten hyphenspezifischen Protein in spezifischer Beziehung steht. Ein erfindungsgemäß verwendeter Antikörper kann auch gegen Strukturen gerichtet sein, die als posttranslationelle Modifikation der hyphenspezifischen Proteine anzusehen sind.

Erfindungsgemäß ist vorgesehen, dass der Antikörper-Chip sowohl monoclonale als auch polyclonale Antikörper enthält, die in der Lage sind, spezifisch eine Struktur eines erfindungsgemäßen

hyphenspezifischen Proteins zu identifizieren und gegebenenfalls zu binden.

Die auf dem erfindungsgemäßen Antikörper-Chip enthaltenen monoclonalen und polyclonalen Antikörper können unter Verwendung von auf dem Fachgebiet wohlbekannten Verfahren hergestellt und isoliert werden. Die Verfahren zur Herstellung monoclonaler Antikörper unter Verwendung der Hybridom-Technologie sind beispielsweise in Schreyer et al., „Hybridoma Techniques“ (1980) oder in den US-Patenten Nr. 4,341,761, Nr. 4,399,121, Nr. 4,472,500 beschrieben.

Als Trägermaterial zur Immobilisierung der Antikörper können die vorstehend für Protein-Chips genannten Materialien, wie Glas, Siliciumdioxid, Nylon, Acrylamid-Derivate, Silica-Materialien, Acopolymere Materialien, wie Fluorpolymere, Metalloxide, usw. verwendet werden. Diese Trägermaterialien werden vorzugsweise vor Immobilisierung der Antikörper vorbehandelt, beispielsweise mit Silan-Kupplungsmitteln. In einer bevorzugten Ausführungsform der Erfindung werden für die Antikörper-Chips die von Joos et al. beschriebenen Epoxy-modifizierte Träger oder Membranen (Joos et al., 2000, Electrophoresis, 21, 2641-2650) verwendet.

Erfindungsgemäß ist vorgesehen, dass die Bindung der gegen hyphenspezifische Proteine gerichteten Antikörper an den Träger über eine chemische oder photochemische Reaktion oder über elektrostatische Wechselwirkungen erfolgt. Wie vorstehend für die Protein-Chips beschrieben, können die erfindungsge-

mäß verwendeten Antikörper beispielsweise mit Hilfe einer Vielzahl üblicherweise verwendeter funktioneller Gruppen und/oder Spacer oder chemischer Vernetzungsmittel, wie bifunktionaler Vernetzungsmittel, am Trägermaterial gebunden und immobilisiert werden.

Eine weitere bevorzugte Ausführungsform der vorliegenden Erfindung betrifft daher Verfahren zur Herstellung von Antikörper-Chips, umfassend die gentechnische Herstellung, Isolierung oder Synthese mindestens eines gegen ein hyphenspezifisches Candida-Protein gerichteten Antikörpers oder Fragments davon, die Modifikation des Antikörpers oder Fragments während oder nach der Isolierung, Synthese oder Herstellung durch den Einbau funktioneller Gruppen oder Spacer-Einheiten, das Auftragen einer wässrigen Lösung des isolierten oder synthetisierten Antikörpers auf ein festes Trägermaterial und die Immobilisierung des Antikörpers an dem Träger mittels chemischer oder photochemischer Reaktion oder elektrostatischer Wechselwirkung.

Eine weitere besonders bevorzugte Ausführungsform der Erfindung betrifft einen Antikörper-Chip, umfassend einen festen Träger und mindestens einen daran fixierten Antikörper, der spezifisch gegen einen Antikörper gegen ein Protein mit der in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 dargestellten Aminosäuresequenz gerichtet ist. Ein derartiger Antikörper-Chip umfasst also Antikörper, die gegen die vorgenannten Antikörper gerichtet sind und diese spezifisch erkennen und binden können. Die Verwendung eines derartigen Antikörpers ermöglicht al-

so den Nachweis von Antikörpern gegen hyphenspezifische Proteine von Candida im Blut, der Lymphe, in Körpersekreten oder anderen Körperflüssigkeiten eines Organismus und somit den Nachweis einer systemischen Candida-Infektion in diesem Organismus, die zur Bildung der in den Proben enthaltenen Antikörper geführt hat.

Die vorliegende Erfindung betrifft auch eine diagnostische Zusammensetzung, umfassend mindestens einen erfindungsgemäßen Nucleotid-Chip, einen erfindungsgemäßen Protein-Chip und/oder einen erfindungsgemäßen Antikörper-Chip. Die Erfindung umfasst daher auch Diagnostikkits, die die erfindungsgemäßen Biochips, das heißt Nucleotid-Chips, Protein-Chips und Antikörper-Chips, geeignete Puffersysteme und geeignete Markierungssysteme enthalten.

Die Erfindung betrifft in einer besonders bevorzugten Ausführungsform Verfahren zur Diagnose von durch Candida-Arten verursachten Krankheiten, insbesondere durch Candida albicans verursachte Krankheiten, wobei eine zu testende Probe in einem geeigneten Medium mit einem erfindungsgemäßen Nucleotid-Chip, einem erfindungsgemäßen Protein-Chip und/oder einem erfindungsgemäßen Antikörper-Chip in Kontakt gebracht wird und eine Interaktion zwischen der zu testenden Probe und mindestens einem der genannten Bio-Chips nachgewiesen wird. Die erfindungsgemäßen Verfahren zur Diagnose von Candida-Erkrankungen basieren auf dem Nachweis der Anwesenheit von Nucleotidsequenzen, Proteinen, Antikörpern oder Fragmenten davon, die im Zusammenhang mit den erfindungsgemäßen hyphenspezifischen Proteinen ste-

hen, in Proben, wie Haut- oder Schleimhautabstrichen, Organbiopsien, Körpersekreten, Blut, Lymphe oder anderen Körperflüssigkeiten etc., die pilzlichen Befall zeigen oder im Verdacht stehen, infiziert zu sein.

Im Zusammenhang mit der vorliegenden Erfindung werden unter „durch Candida-Arten verursachten Krankheiten, Krankheitszuständen oder Infektionen“ solche Erkrankungen verstanden, die ausschließlich von Candida-Arten, insbesondere jedoch von *Candida albicans* verursacht werden. Der Begriff umfasst daher auch alle Erkrankungen, die von Candida-Arten wie *C. tropicalis*, *C. krusei*, *C. parapsilosis* und *C. guilliermondii*, oder *Torulopsis (Candida) glabrata* verursacht werden. Erfindungsgemäß werden unter diesem Begriff auch Krankheiten oder Krankheitszustände verstanden, die primär andere Ursachen haben und bei denen die Candida-Arten am Gesamtkrankheitsbild nur beteiligt sind beziehungsweise zusätzliche Symptome hinzufügen, zum Beispiel opportunistische Infektionen. Der Begriff „durch Candida-Arten verursachten Krankheiten, Krankheitszuständen oder Infektionen“ umfasst insbesondere solche Erkrankungen wie Candida-Mykosen oder Candidosen, die sich im wesentlichen in drei Hauptformen unterteilen lassen. Die erste Hauptform von Candidosen ist durch eine saprophytäre Besiedlung der Haut und Schleimhäute, insbesondere in den äußeren Genitalien, im Mund, Nase-Rachenraum und im Verdauungstrakt gekennzeichnet. Die zweite Candidosen-Hauptform umfasst Infektionen von Haut und Schleimhäuten und wird in erheblichem Maße durch Faktoren wie Schwangerschaft, Diabetes mellitus, schwere Er-

krankungen oder Traumen, Cytostatika- und Antibiotikatherapie sowie Alkoholkrankheit begünstigt. Die dritte Hauptform umfasst tiefe Organmykosen bei immunsupprimierten Patienten mit zellulärer Immunschwäche, insbesondere im Bereich der Atemwege, seltener als Candida-Endocarditis, Candida-Meningitis, Candida-Nephritis und Candida-Endophthalmitis.

Aufgrund der Hyphenspezifität und der damit verbundenen Korrelation zu einem Candida-verursachten Krankheitsbild weist die Anwesenheit von Nucleotidsequenzen, Proteinen, Antikörpern oder Fragmenten davon, die im Zusammenhang mit den erfindungsgemäßen hyphenspezifischen Proteinen stehen, auf eine Candida-verursachte Krankheit hin. Der Nachweis der vorgenannten Substanzen erfolgt durch Bindung und/oder Hybridisierung an mindestens einen der erfindungsgemäßen Biochips, die die nachzuweisenden Substanzen spezifisch erkennen.

Sollen im Zusammenhang mit den hyphenspezifischen Proteinen stehende Nucleotidsequenzen in einer Probe nachgewiesen werden, erfolgt deren Nachweis durch Hybridisierung mit dem erfindungsgemäßen Nucleotid-Chip. Dazu werden Nucleinsäuren, beispielsweise DNA oder mRNA, aus einer Probe oder aus einer eigens angelegten Kultur isoliert und/oder amplifiziert, beispielsweise mittels PCR-Verfahren. Die extrahierten Nucleinsäuren werden anschließend markiert, zum Beispiel mit Fluoreszenzfarbstoffen, Enzymen oder radioaktiven Gruppen. Hybridisiert die extrahierte und markierte DNA mit dem Nucleotid-Chip, so zeigt dies die Anwesenheit von Candida in



der untersuchten Probe. Hybridisiert die extrahierte mRNA mit dem erfindungsgemäßen Nucleotid-Chip, so weist dies darauf hin, dass die Probe hyphal wachsende Candida-Formen enthält.

Sollen hyphenspezifische Candida-Proteine in einer Untersuchungsprobe nachgewiesen werden, werden Proteine aus der Probe extrahiert und markiert und anschließend mit dem erfindungsgemäßen Antikörper-Chip, der Antikörper gegen hyphenspezifische Proteine enthält, inkubiert. Die Bindung mindestens eines markierten Proteins am erfindungsgemäßen Antikörper-Chip weist auf die Anwesenheit hyphal wachsender Candida-Zellen hin. Der Nachweis von Antikörpern gegen Candida-Proteine in Körperflüssigkeiten kann sowohl unter Verwendung des erfindungsgemäßen Protein-Chips als auch unter Verwendung des erfindungsgemäßen Antikörper-Chips, der Antikörper enthält, die gegen die Antikörper gegen hyphenspezifische Proteine gerichtet sind, erfolgen.

Die vorliegende Erfindung betrifft auch Verfahren zum Auffinden und identifizieren von Substanzen, die gegen Candida-verursachte Krankheiten therapeutisch wirksam sind, wobei eine zu testende Substanz in einem geeigneten Medium mit einem erfindungsgemäßen Nucleotid-Chip, einem erfindungsgemäßen Protein-Chip oder einem erfindungsgemäßen Antikörper-Chip in Kontakt gebracht und eine Wechselwirkung zwischen der zu testenden Substanz und mindestens einem der genannten Chips nachgewiesen wird. So können die erfindungsgemäßen Nucleotid- und Protein-Chips, wie vorstehend beschrieben, verwendet werden, um Substanzen, beispielsweise Proteine, zu

identifizieren, die in vivo an Nucleotidsequenzen, die hyphenspezifisch exprimierte Proteine codieren beziehungsweise die Expression dieser Proteine regulieren, oder an die hyphenspezifischen Proteine selbst binden. Solche bindenden Substanzen, insbesondere Proteine, können potentiell als Medikamente gegen Candida-verursachte Krankheiten geeignet sein, wenn sie beispielsweise in der Lage sind, durch Bindung an regulatorische Nucleotidsequenzen die Transkription hyphenspezifischer Proteine zu hemmen oder zu unterbinden oder wenn sie durch Bindung an die hyphenspezifischen Proteine deren Aktivität hemmen oder unterbinden können. Wenn solche Substanzen, die an erfindungsgemäße Nucleotid- oder Proteinchips binden, die Transkription der hyphenspezifischen Proteine induzieren oder fördern oder die Aktivität hyphenspezifischer Proteine derartiger Proteine begünstigen, können die erfindungsgemäßen Biochips verwendet werden, um weitere Substanzen zu identifizieren, die die Interaktion zwischen einem hyphenspezifischen Protein beziehungsweise der dieses codierenden Nucleotidsequenz und der daran bindenden Substanz, insbesondere eines bindenden Proteins, beeinflussen oder hemmen können. Auch solche Substanzen sind potentiell als Medikamente zur Behandlung Candida-verursachter Krankheiten geeignet.

Die vorliegende Erfindung wird durch das folgende Sequenzprotokoll und die folgenden Figuren und Beispiele erläutert.

Das Sequenzprotokoll ist Teil dieser Beschreibung und enthält die Sequenzprotokolle SEQ ID Nr. 1 bis

18. Jede der nachstehend aufgeführten Aminosäuresequenzen wurde aus der entsprechenden DNA-Sequenz abgeleitet und dann teilweise durch Sequenzierung der isolierten Proteine verifiziert.

SEQ ID Nr. 1 stellt die codierende Cap33a DNA-Sequenz aus Contig4-2149 dar.

SEQ ID Nr. 2 stellt die codierende Cap33b DNA-Sequenz aus Contig4-2501 dar.

SEQ ID Nr. 3 stellt die codierende Cap18p DNA-Sequenz aus Contig4-2069 dar.

SEQ ID Nr. 4 stellt die codierende Cap19p DNA-Sequenz aus Contig4-2069 dar.

SEQ ID Nr. 5 stellt die Aminosäuresequenz von Cap33a dar.

SEQ ID Nr. 6 stellt die Aminosäuresequenz von Cap33b dar.

SEQ ID Nr. 7 stellt die Aminosäuresequenz von Cap18p dar.

SEQ ID Nr. 8 stellt die Aminosäuresequenz von Cap19p dar.

SEQ ID Nr. 9 stellt die gesamte DNA-Sequenz von Contig4-2149 dar.

SEQ ID Nr. 10 stellt die gesamte DNA-Sequenz von Contig4-2501 dar.

SEQ ID Nr. 11 stellt die gesamte DNA-Sequenz von Contig4-2069 dar.

SEQ ID Nr. 12 stellt den Promoter-Bereich von Cap18p und Cap19p dar.

SEQ ID Nr. 13 stellt die Cap15p codierende DNA-Sequenz aus Contig5-3226 dar.

SEQ ID Nr. 14 stellt die Aminosäuresequenz von Cap15p dar. Bei Cap15p handelt es sich wahrscheinlich um eine Nucleosid-Diphosphat-Kinase.

SEQ ID Nr. 15 stellt die Cap20p codierende DNA-Sequenz aus Contig4-2178 dar.

SEQ ID Nr. 16 stellt die Aminosäuresequenz von Cap20p dar. Bei Cap20p handelt es sich wahrscheinlich um eine Glutathionperoxidase.

SEQ ID Nr. 17 stellt die Cap40p codierende DNA-Sequenz aus Contig5-2806 dar.

SEQ ID Nr. 18 stellt die Aminosäuresequenz von Cap40p dar. Bei Cap40p handelt es sich wahrscheinlich um eine Fructose-Biphosphat-Aldolase.

Figur 1 zeigt im linken Teil mikroskopische Aufnahmen des virulenten, hyphal wachsenden *Candida albicans*-Stammes Sc5315 und des avirulenten, hefeartig wachsenden *Candida*-Stammes Can34 ( $\Delta$ cph1 $\Delta$ efg1). Aus beiden Stämmen wurde RNA isoliert und in cDNA umgeschrieben. Die markierte cDNA wurde mit einem erfindungsgemäßen Nucleotid-Chip, auf dem CAP33, CAP19 und CAP18 codierende Nucleotidsequenzen fi-

xiert waren, hybridisiert. Die Ergebnisse dieser Hybridisierung sind im rechten Teil von Figur 1 zu sehen. Während die cDNA aus dem avirulenten Stamm mit keiner der auf dem Nucleotid-Chip enthaltenen Nucleotidsequenzen hybridisierte, zeigte die cDNA des virulenten Stammes mit allen drei immobilisierten Nucleotidsequenzen starke Hybridisierungssignale. Wird der *C. albicans*-Stamm Sc5315 jedoch unter Bedingungen kultiviert, unter denen keine Hyphen entstehen, wird dasselbe Ergebnis wie für den avirulenten Stamm Can34 ( $\Delta$ cph1 $\Delta$ efg1) erhalten.

Figur 2a zeigt mikroskopische Aufnahmen des virulenten, hyphal wachsenden *Candida albicans*-Stammes Sc5315, und des avirulenten, hefeartig wachsenden *Candida*-Stammes Can34 ( $\Delta$ cph1 $\Delta$ efg1). Darunter sind Abbildungen differentieller Proteom-Analysen der beiden Stämme nach Kultivierung in  $\alpha$ -MEM-Medium gezeigt. Aus diesen Abbildungen geht hervor, dass Sc5314 bei Kultivierung in  $\alpha$ -MEM-Medium die Proteine p33a und p33b exprimiert, Can34 ( $\Delta$ cph1 $\Delta$ efg1) jedoch nicht.

Figur 2b zeigt die Ergebnisse einer Northern-Blot-Analyse unter Verwendung von RNAs aus diesen beiden Stämmen, die zuvor entweder in YPD-Medium oder in  $\alpha$ -MEM-Medium kultiviert worden waren. Ebenfalls einbezogen wurden RNAs aus dem virulenten Stamm Can16 ( $\Delta$ cph1) und dem Stamm Can33 ( $\Delta$ efg1), der keine Hyphenbildung zeigt und eine stark verminderte Virulenz besitzt, wobei beide Stämme vor RNA-Extraktion in  $\alpha$ -MEM-Medium kultiviert worden waren. Bei Hybridisierung mit cap18-, cap-19 und cap33-spezifischen Sonden wurden Hybridisierungssignale

mit den aus in  $\alpha$ -MEM-Medium kultivierten Stämmen Sc5314 und Can16 isolierten RNAs gefunden. Die aus dem in YPD-Medium kultivierten Stamm Sc5314 isolierte RNA zeigte hingegen keine Hybridisierungssignale. Die RNA aus dem avirulenten Stamm Can34 (kultiviert entweder in YPD- oder in  $\alpha$ -MEM-Medium) zeigte ebenfalls keine Hybridisierungssignale. Das gleiche Ergebnis wurde auch bei der RNA erhalten, die aus dem in  $\alpha$ -MEM-Medium kultivierten Stamm Can33 isoliert worden war. Als Kontrolle wurde eine Aktin-spezifische Sonde (ACT1) eingesetzt.

Figur 3 zeigt die Ergebnisse von 2D-Gelelektrophoresen von Proteinextrakten aus den in  $\alpha$ -MEM-Medium kultivierten Stämmen Sc5315 und Can34 ( $\Delta$ cph1 $\Delta$ efg1). Während Sc5315 die hyphenspezifischen Proteine p33a, p33b, p40, p15, p18, p19 und p20 exprimiert, werden diese Proteine in Can34 ( $\Delta$ cph1 $\Delta$ efg1) nicht exprimiert.

Weitere vorteilhafte Ausgestaltungen der Erfindung sind in den Unteransprüchen beschrieben.

Die Erfindung wird anhand des folgenden Beispiels näher beschrieben.

#### **Beispiel 1:**

##### *Die Isolierung der Proteine*

Die Proteine wurden aus dem klinischen Isolat Sc5314 durch differenzielle 2D-Gelelektrophorese wie folgt isoliert:

Zur Isolierung der Proteine wurden der virulente *Candida albicans*-Stamm Sc5314 und der avirulente *Candida albicans* Stamm Can34 (HLC69) (Lo et al., 1997, Cell, 90, 939-949) gleichzeitig in Vollmedium (YPD: 20g/l Bacto-Pepton; 10g/l Hefe-Extrakt; 0,15g/l L-Tryptophan) über Nacht angezogen, in  $\alpha$ -MEM Medium (#22571 Life Technologies/Gibco) mit 2 % Glucose inokuliert (1:100) und für 24 h bei 37°C auf einem Rundschüttler inkubiert. Die so gewonnenen Zellen wurden pelletiert und in einem nicht detergenzhaltigen salinen Puffer (PGSK-Puffer: 0,52g/l  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ ; 8,8g/l  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ ; 2,8g/l NaCl; 0,372g/l KCl; 11g/l Glucose) mit Glasperlen aufgeschlossen. Der daraus isolierte Proteinextrakt wurde mittels isoelektrischer Fokussierung und danach mittels SDS-PAGE aufgetrennt. Die Gele wurden mit Silber (vgl. Figur 2a) oder Coomassie (vgl. Figur 2b) angefärbt. Die Proteinspots, die nur in einem der beiden Gele sichtbar waren, wurden aus den coomassiegefärbten Gelen ausgestochen und deren Sequenz bestimmt. Es konnte gezeigt werden, dass die identifizierten Proteine ausschließlich in Sc5314 in  $\alpha$ -MEM gebildet werden (Figuren 2a und 3).

Aufgrund der Aminosäure-Sequenz, die durch Edmann-Abbau von tryptischen Fragmenten des Proteins eindeutig zu bestimmen war, konnte die zugehörige DNA-Sequenz über Datenbankvergleiche identifiziert werden. Die DNA-Sequenzen sowie die flankierenden Bereiche wurden durch PCR aus genomischer DNA von Sc5314 amplifiziert und cloniert. Weiterhin wurde die entsprechende DNA-Sequenz aus genomischen Bibliotheken (Liu et al., 1995, Science, 266, 1723-

1726) mittels Hybridisierung der durch PCR gewonnenen radioaktiv markierter Fragmente isoliert.

Die codierenden Sequenzen zu jedem der sieben identifizierten Proteine wurden aus den clonierten PCR-Fragmenten -mittels PCR- entfernt und durch Selektionsmarker (URA3) ersetzt (Fonzi and Irwin, 1993; Genetics, 134, 717-728). Diese Konstrukte werden zur Deletion der codierenden Sequenz in *C. albicans* verwendet. Weiterhin wurden die offenen Leserahmen zu allen sieben Proteinen sowie die Terminationssequenzen mittels PCR isoliert und in Vektoren mit regulierbaren PCK1- und MET3-Promotoren (Leuker et al., Gene, 1997, 19, 192(2), 235-40; Care et al., Mol Microbiol., 1999, 34(4), 792-8.) zur Expression in *C. albicans* beziehungsweise mit regulierbaren GAL1-10- und MET25-Promotoren (Mumberg et al., Nucleic Acids Res. 1994, 22 (25), 5767-8.) zur Expression in *S. cerevisiae* und in geeignete Vektoren (pMAL, PGEx etc.) zur Expression der Proteine in Bakterien cloniert.

#### **Beispiel 2:**

##### *Nachweis der hyphenspezifischen Expression der erfindungsgemäßen Proteine*

Die Regulation der hyphenspezifisch exprimierten Proteine findet auf Transkriptionsebene statt, da die mRNA zu allen sieben Proteinen ausschließlich in  $\alpha$ -MEM gewachsenen Sc5314-Kulturen nachweisbar ist, nicht jedoch in Sc5314, der in Vollmedium kultiviert wurde, oder im avirulenten Stamm Can34



( $\Delta$ cph1 $\Delta$ efg1), der in Vollmedium oder  $\alpha$ -MEM-Medium kultiviert wurde.

Figur 2b zeigt als Beispiel eine Northern-Analyse von RNA aus den in Vollmedium (YPD oder  $\alpha$ -MEM) kultivierten Stämmen Sc5314 und Can34 ( $\Delta$ cph1 $\Delta$ efg1). Bei dieser Northern-Analyse wurde zudem RNA der in  $\alpha$ -MEM-Medium kultivierten Stämme Can16 ( $\Delta$ cph1) und Can33 ( $\Delta$ efg1) (Lo et al., 1997, Cell, 90, 939-949) aufgetragen. Can16 ( $\Delta$ cph1) wurde als Stamm beschrieben, der eine mit Sc5314 vergleichbare Virulenz aufweist und Hyphenbildung zeigt. Can33 ( $\Delta$ efg1) zeigt hingegen keine Hyphenbildung und seine Virulenz ist stark reduziert (Lo et al., 1997, Cell, 90, 939-949).

Im virulenten Stamm Sc5314, der in  $\alpha$ -MEM-Medium kultiviert wurde, können mit Cap33-, Cap18- und Cap19-spezifischen Sonden die entsprechenden mRNAs nachgewiesen werden. Hingegen lassen sich in Sc5314, der in YPD-Medium kultiviert wurde, mit den vorstehend genannten Sonden keine entsprechenden mRNAs nachweisen. Im virulenten Stamm Can16, kultiviert in  $\alpha$ -MEM-Medium, können unter Verwendung der Sonden die entsprechenden mRNAs ebenfalls nachgewiesen werden. Der avirulente Stamm Can34 ( $\Delta$ cph1 $\Delta$ efg1), kultiviert entweder in YPD-Medium oder in  $\alpha$ -MEM-Medium, enthält keine entsprechende mRNAs. Auch der in  $\alpha$ -MEM-Medium kultivierte Stamm Can33 enthält keine entsprechenden mRNAs.

Die Ergebnisse dieser Northern-Blot-Analyse zeigen einerseits die hyphenspezifische Expression der Proteine mit SEQ ID Nr. 5, 5, 7, 8, 14, 16 und 18

in den virulenten Stämmen Sc5314 und Can16. Andererseits zeigen diese Ergebnisse, dass die hyphenspezifische Expression dieser Proteine hauptsächlich auf der Transkriptionsebene reguliert wird. Dies ist für die Verwendung als diagnostisches Mittel bedeutsam, da die Anwesenheit von mRNA, die im Zusammenhang mit einem der vorstehend genannten hyphenspezifischen Proteinen steht, in einer zu testenden Probe ein Hinweis auf das Vorkommen hyphal wachsender virulenter Candida-Formen in dieser Probe ist.

Proteinextrakte aus den Stämmen Sc5314 und Can34 ( $\Delta$ cph1 $\Delta$ efg1), die in  $\alpha$ -MEM-Medium kultiviert worden waren, wurden entsprechend üblichen Verfahren einer 2D-Gelelektrophorese unterworfen. Wie in Figur 3 zu sehen, sind die hyphenspezifisch exprimierten Proteine p15, p18, p19, p20, p33a, p33b und p40 ausschließlich im hyphal wachsenden Stamm Sc5314 nachweisbar, nicht jedoch im avirulenten Stamm Can34 ( $\Delta$ cph1 $\Delta$ efg1). In Figur 2a ist darüber hinaus dargestellt, dass der in  $\alpha$ -MEM-Medium kultivierte Stamm Sc5314 die Proteine Cap33a und Cap33b produziert, der Stamm Can34 jedoch nicht.

### Beispiel 3:

#### *Nachweis von hyphenspezifische Proteine codierenden Nucleotidsequenzen mittels eines Nucleotid-Chips*

Auf einer Poly-L-Lysin-beschichteten Trägeroberfläche wurden Nucleotidsequenzen, die die hyphenspezifischen Proteine CAP33, CAP19 und CAP18 codieren, fixiert. Der so erhaltene Nucleotid-Chip wurde un-

-42-

ter üblichen Hybridisierungsbedingungen mit cDNAs hybridisiert, die aus mRNA des hyphal wachsenden *Candida albicans*-Stammes Sc5315 und des avirulenten, hefeartig wachsenden *Candida*-Stammes Can34 ( $\Delta$ cph1 $\Delta$ efg1) hergestellt worden waren. Die Ergebnisse dieser Hybridisierung sind in Figur 1 dargestellt. Während die cDNA aus dem avirulenten Stamm mit keiner der auf dem Nucleotid-Chip enthaltenen Nucleotidsequenzen hybridisierte, zeigte die cDNA des virulenten Stammes mit allen drei immobilisierten Nucleotidsequenzen starke Hybridisierungssignale. Dieser Versuch zeigt, dass unter Verwendung eines erfindungsgemäßen Nucleotid-Chips spezifische Nucleinsäuren nachgewiesen werden können und dass damit hyphal wachsende virulente *C. albicans*-Stämme von nicht-hyphal wachsenden, avirulenten *C. albicans*-Stämmen unterschieden werden können.

**Ansprüche**

1. Nucleotid-Chip, umfassend einen festen Träger und mindestens eine daran fixierte Nucleotidsequenz, welche für die Identifizierung und Transkription eines ein hyphenspezifisches Proteincodierenden Gens geeignet ist, ausgewählt aus der Gruppe bestehend aus:

(a) einer Nucleotidsequenz, definiert in SEQ ID Nr. 1, 2, 3, 4, 12, 13, 15 oder 17, oder eines komplementären Strangs oder Teils davon,

(b) einer Nucleotidsequenz, codierend die Aminosäuresequenz, definiert in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18, oder eines komplementären Strangs oder Teils davon und

(c) einer Nucleotidsequenz, die mit einer der (a) oder (b) genannten Nucleotidsequenzen hybridisiert.

2. Nucleotid-Chip nach Anspruch 1, wobei die Nucleotidsequenz eine DNA-, RNA- oder PNA-Sequenz ist.

3. Protein-Chip, umfassend einen festen Träger und mindestens ein daran fixiertes hyphenspezifisches Protein, ausgewählt aus der Gruppe bestehend aus:

-44-

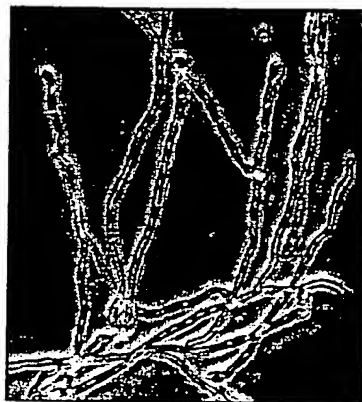
- (a) einem Protein mit einer Aminosäuresequenz, definiert in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18, oder einem Fragment davon, und
  - (b) einem Protein mit einer Aminosäuresequenz, die zu einer der in SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 definierten Aminosäuresequenz eine Sequenzidentität von mindestens 80% aufweist, oder einem Fragment davon.
4. Protein-Chip nach Anspruch 3, wobei das Fragment des Proteins ein eine Antigen-Determinante umfassender Protein-Bereich ist.
  5. Antikörper-Chip, umfassend einen festen Träger und mindestens einen daran fixierten Antikörper, der spezifisch gegen ein Protein mit der SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 gerichtet ist.
  6. Antikörper-Chip nach Anspruch 5, wobei der Antikörper ein monoklonaler, polyclonaler und/oder modifizierter Antikörper ist.
  7. Antikörper-Chip, umfassend einen festen Träger und mindestens einen daran fixierten Antikörper, der spezifisch gegen einen Antikörper gegen ein Protein mit der SEQ ID Nr. 5, 6, 7, 8, 14, 16 oder 18 gerichtet ist.
  8. Diagnostische Zusammensetzung, enthaltend einen Nucleotid-Chip nach Anspruch 1 oder 2, einen Protein-Chip nach Anspruch 3 oder 4 oder einen Antikörper-Chip nach einem der Ansprüche 5 bis 7.

9. Verfahren zum Diagnostizieren einer durch Arten der Gattung Candida verursachten Krankheit, wobei eine zu testende Probe in einem geeigneten Medium mit einem Nucleotid-Chip nach Anspruch 1 oder 2, einen Protein-Chip nach Anspruch 3 oder 4 oder einen Antikörper-Chip nach einem der Ansprüche 5 bis 7 in Kontakt gebracht wird und eine Interaktion zwischen der zu testenden Probe und mindestens einem der genannten Chips nachgewiesen wird.

10. Verfahren zum Auffinden und Identifizieren therapeutisch gegen durch Arten der Gattung Candida verursachten Krankheiten wirksamer Substanzen, wobei eine zu testende Substanz in einem geeigneten Medium mit einem Nucleotid-Chip nach Anspruch 1 oder 2, einen Protein-Chip nach Anspruch 3 oder 4 oder einen Antikörper-Chip nach einem der Ansprüche 5 bis 7 in Kontakt gebracht wird und eine Interaktion zwischen der zu testenden Substanz und mindestens einem der genannten Chips nachgewiesen wird.

# Abb. 1

Sc5315  
virulent



→ RNA → cDNA →

CHIP

Sc5315  
virulent



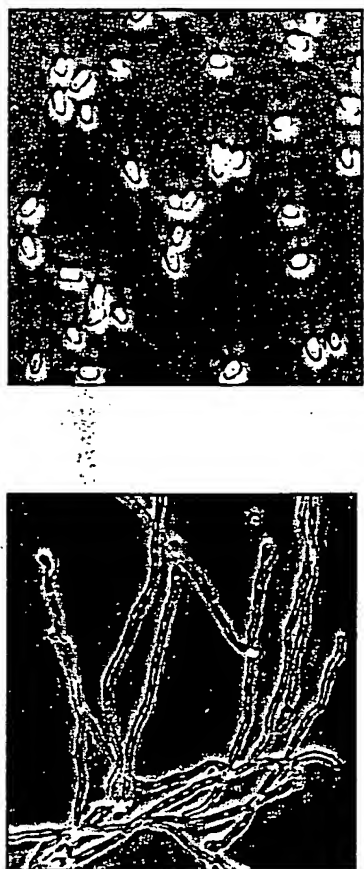
→ RNA → cDNA →

Can34  
avirulent

CAP33 CAP19 CAP18

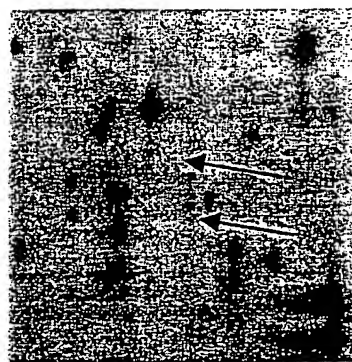
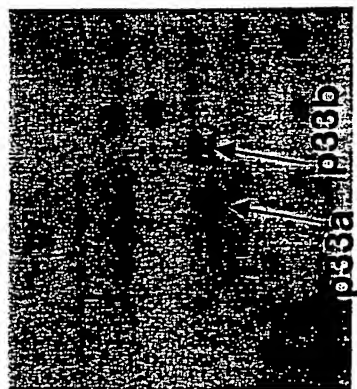
Can34 ( $\Delta cph1/\Delta efg1$ )  
avirulent

# Abb. 2 a)



Sc5314

Can34 ( $\Delta$ efg1/ $\Delta$ cph1)



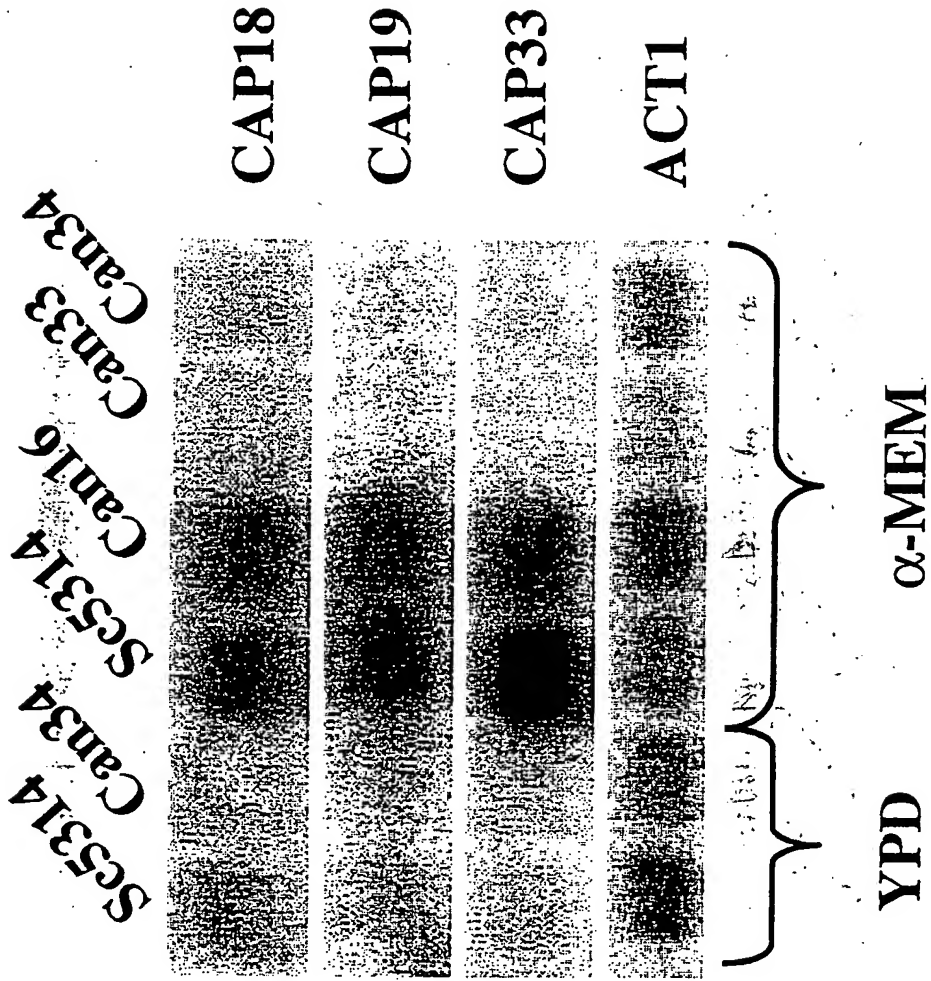
Differentielle Proteom-Analyse, Kulturmedium  $\alpha$ -MEM



# Abb. 2

b)

Northern  
blott

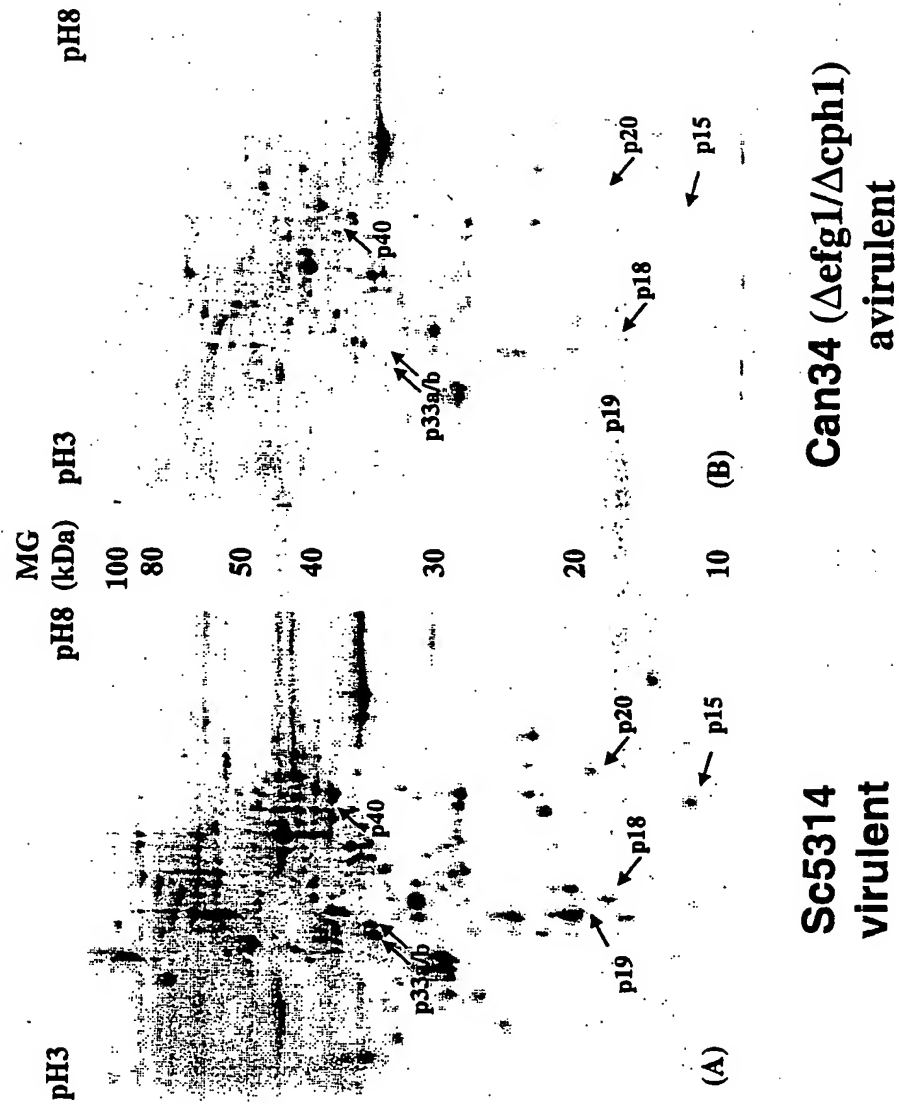


Kulturmedium:

YPD

α-MEM

# Abb. 3



## SEQUENZPROTOKOLL

&lt;110&gt; Fraunhofer-Gesellschaft zur Förderung.....

&lt;120&gt; Hyphenspezifische Faktoren aus Candida albicans

&lt;130&gt; 23874

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 18

&lt;170&gt; PatentIn Ver. 2:1

&lt;210&gt; 1

&lt;211&gt; 900

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 1

```

atgtcctaaag tctcaattac tatcatcggg ttgaatgggt tcttaggtaa accagttctt 60
gaagctatca attctgggtat ttttgatgat aagatcaact tcccaatcaa ggcaattacc 120
agaaaaggaac cagaaactaa gaatgacaaa attgaatatg ttgtttctga aatcaatgaa 180
gaatcaatta aatcaacttt gagccaaaaa ttatctggta ctgatgttat tattgaatta 240
attggtccaa atccagaggc tttcgccaat atcgaaaatt tagttgatgc aattaaacca 300
aaattattta tcccatcaca atttggtact gatattccta aagttgatga atatgctcca 360
gggttttttag gaattaaaac tcaacattca gaaaatgtca gaaaatcagg agttaaagtt 420
gttgatatta taacttcgtt atttgctgtt ccaggagctt ttctttatga atgggttggg 480
tcaactggta ttaatgctga agacagaact gttaaactca ttggtgacat taatcaacaa 540
tttgatattt ctaaattaga agatgttggg aaagctgtac ttctattgc tactaatcct 600
aatccaagag aattaccaga taccattaga attggttctg atagaattac tgttaaagat 660
gtaattgata gatactctaa agatcataat gttgaattga aaattgtttc agaacaatct 720
gcagaagacg ccaagaaaga gtttactgaa tctttgaaag ttggttttga tgggtgataaa 780
ttcttatggg atttacaagt tattgctgct caagggttag ataaagggtt actctccagt 840
aaattggata atgaattggg taatccaggg gaatctttat ggaaatgggg caagtactaa 900

```

&lt;210&gt; 2

&lt;211&gt; 900

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 2

```

atgtcctaaag tctcaattac tatcatcggg ttgaatgggt tcttaggtaa accagttctt 60
gaagctatca attctgggtat ttttgacgat aaaatcaatt tcccaattaa agctattaca 120
agaaaagaac cggaaaactaa aaatgacaaa attgaatagc ttgtttctga aatcaatgaa 180
gaatcaatta aatcaacttt gagccaaaaa ttatccggta ctgatgttat tattgaatta 240
attggtccaa atccagaggc tttcgcta atcgaaaaat taattgatgc aattaaacca 300
aaattattca tcccatcaca atttggtact gatattccta aagttgatga atatgctcca 360
gggttttttag gaatcaaaaac tcaacattca gaaaatgtca gaaaattagg agttaaagtt 420
gttgatatta taacttcgtt atttgctgtt ccaggagctt ttctttatga atgggttggg 480
tcaactggta tcaatgctga tgacaaaact gttaaactta ttggtgacat taatcaacaa 540
tttgatattt ctaaattaga agatgttggg aaagctgtac ttctattgc tactaatcct 600
aatccaagag aattaccaga taccattaga attggttctg atagaattac tgtcaaagat 660
gtcattgata gatattctaa agatcataat gttgaattga aagttgtttc tgaacaatct 720
gcagaagatg ccaagaaaga gtttactgaa tctttgaaag ctggttttga tgggtgagaaa 780
ttcttatggg atttacaagt tattgctgct caagggttag ataaagggtt actctccagt 840
aaattggaca acgaattggg caaccaggg gagtctttat ggaaatgggg caagtactaa 900

```

&lt;210&gt; 3

&lt;211&gt; 501

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 3

```

atggcctcct cagtaaagtt ggctacggca cttaaacaac gtgctatatt gacaaaagaa 60
ttgtctgaat tagatgataa aatacaatct tcattgattc tgcaagttgg tatgaaaaaa 120
atcaatgac cagataaatt gtatttagat tatgttgcta aatctcaaga attggctaaa 180
ttggtatcat caataaatta tactaataat ataactcaa ttgaacttga tttgacaatg 240
ggaaagtatg ataatactat aaaaacaatt aatgatgcat taatttgtcg agaccgaata 300
tttaaaaaat tacaatttgt gaaaaaata tcaacagcag gtaaagaaca accattagat 360
tccaaagatg aaattaaatt tgtatcattt attgatgttg ataaatatga tactttggcc 420
caagaattaa atactcaatt tgagaatttg aatttgaaat tacaagaaat aaattggcaa 480
gttgatcttg ttgatata a                                     501

```

&lt;210&gt; 4

&lt;211&gt; 486

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 4

```

atgaaattag ctgaagcatt aaatttaaaa aagaacttgg aaagagatgc tgggtgaactt 60
aaatcattaa ttcttaaatg ttgtcaagct caaactggcg aaaaccctcc atttgatcct 120
aatgaattat ttgaacaata tgaagaaatt gataaattaa ttactgatat aactattaaa 180
atacaacgaa ccaataatga aataaagttt gcctatgata atgataataa gtctaataaa 240
gaaccacttc gatcaatgac acaagctatt gctgatattg atgatttaga aagaacaaatc 300
aatgtgacag atgatataat tcataatggg attattacaa aactgtattc gaccaagaag 360
attgctgatg tgtcacatgt tgacgtgggt gcatatgaca agacaagaaa gaaaatgaat 420
gagagattag ataaattaaa acttcgtata cagtcggcaa attgggaatt tgatctaatt 480
gattaa                                     486

```

&lt;210&gt; 5

&lt;211&gt; 299

&lt;212&gt; PRT

&lt;213&gt; Candida albicans

&lt;400&gt; 5

```

Met Ser Lys Val Ser Ile Thr Ile Ile Gly Leu Asn Gly Phe Leu Gly
  1           5           10           15

```

```

Lys Pro Val Leu Glu Ala Ile Asn Ser Gly Ile Phe Asp Asp Lys Ile
          20           25           30

```

```

Asn Phe Pro Ile Lys Ala Ile Thr Arg Lys Glu Pro Glu Thr Lys Asn
          35           40           45

```

```

Asp Lys Ile Glu Tyr Val Val Ser Glu Ile Asn Glu Glu Ser Ile Lys
          50           55           60

```

```

Ser Thr Leu Ser Gln Lys Leu Ser Gly Thr Asp Val Ile Ile Glu Leu
          65           70           75           80

```

```

Ile Gly Pro Asn Pro Glu Ala Phe Ala Asn Ile Glu Asn Leu Val Asp
          85           90           95

```

```

Ala Ile Lys Pro Lys Leu Phe Ile Pro Ser Gln Phe Gly Thr Asp Ile
          100          105          110

```

```

Pro Lys Val Asp Glu Tyr Ala Pro Gly Phe Leu Gly Ile Lys Thr Gln
          115          120          125

```

```

His Ser Glu Asn Val Arg Lys Ser Gly Val Lys Val Val Asp Ile Ile

```

130

135

140

Thr Ser Leu Phe Ala Val Pro Gly Ala Phe Leu Tyr Glu Trp Val Gly  
145 150 155 160

Ser Thr Gly Ile Asn Ala Glu Asp Arg Thr Val Lys Leu Ile Gly Asp  
165 170 175

Ile Asn Gln Gln Phe Asp Ile Ser Lys Leu Glu Asp Val Gly Lys Ala  
180 185 190

Val Leu Ser Ile Ala Thr Asn Pro Asn Pro Arg Glu Leu Pro Asp Thr  
195 200 205

Ile Arg Ile Gly Ser Asp Arg Ile Thr Val Lys Asp Val Ile Asp Arg  
210 215 220

Tyr Ser Lys Asp His Asn Val Glu Leu Lys Ile Val Ser Glu Gln Ser  
225 230 235 240

Ala Glu Asp Ala Lys Lys Glu Phe Thr Glu Ser Leu Lys Val Gly Phe  
245 250 255 260

Asp Gly Asp Lys Phe Leu Trp Tyr Leu Gln Val Ile Ala Ala Gln Gly  
260 265 270

Leu Asp Lys Gly Leu Leu Ser Ser Lys Leu Asp Asn Glu Leu Val Asn  
275 280 285

Pro Gly Glu Ser Leu Trp Lys Trp Gly Lys Tyr  
290 295

&lt;210&gt; 6

&lt;211&gt; 299

&lt;212&gt; PRT

&lt;213&gt; Candida albicans

&lt;400&gt; 6

Met Ser Lys Val Ser Ile Thr Ile Ile Gly Leu Asn Gly Phe Leu Gly  
1 5 10 15

Lys Pro Val Leu Glu Ala Ile Asn Ser Gly Ile Phe Asp Asp Lys Ile  
20 25 30

Asn Phe Pro Ile Lys Ala Ile Thr Arg Lys Glu Pro Glu Thr Lys Asn  
35 40 45

Asp Lys Ile Glu Tyr Val Val Ser Glu Ile Asn Glu Glu Ser Ile Lys  
50 55 60

Ser Thr Leu Ser Gln Lys Leu Ser Gly Thr Asp Val Ile Ile Glu Leu  
65 70 75 80

Ile Gly Pro Asn Pro Glu Ala Phe Ala Asn Ile Glu Lys Leu Ile Asp  
85 90 95

Ala Ile Lys Pro Lys Leu Phe Ile Pro Ser Gln Phe Gly Thr Asp Ile  
100 105 110

Pro Lys Val Asp Glu Tyr Ala Pro Gly Phe Leu Gly Ile Lys Thr Gln  
115 120 125

His Ser Glu Asn Val Arg Lys Leu Gly Val Lys Val Val Asp Ile Ile  
 130 135 140  
 Thr Ser Leu Phe Ala Val Pro Gly Ala Phe Leu Tyr Glu Trp Val Gly  
 145 150 155 160  
 Ser Thr Gly Ile Asn Ala Asp Asp Lys Thr Val Lys Leu Ile Gly Asp  
 165 170 175  
 Ile Asn Gln Gln Phe Asp Ile Ser Lys Leu Glu Asp Val Gly Lys Ala  
 180 185 190  
 Val Leu Ser Ile Ala Thr Asn Pro Asn Pro Arg Glu Leu Pro Asp Thr  
 195 200 205  
 Ile Arg Ile Gly Ser Asp Arg Ile Thr Val Lys Asp Val Ile Asp Arg  
 210 215 220  
 Tyr Ser Lys Asp His Asn Val Glu Leu Lys Val Val Ser Glu Gln Ser  
 225 230 235 240  
 Ala Glu Asp Ala Lys Lys Glu Phe Thr Glu Ser Leu Lys Ala Gly Phe  
 245 250 255  
 Asp Gly Glu Lys Phe Leu Trp Tyr Leu Gln Val Ile Ala Ala Gln Gly  
 260 265 270  
 Leu Asp Lys Gly Leu Leu Ser Ser Lys Leu Asp Asn Glu Leu Val Asn  
 275 280 285  
 Pro Gly Glu Ser Leu Trp Lys Trp Gly Lys Tyr  
 290 295

<210> 7  
 <211> 166  
 <212> PRT  
 <213> Candida albicans

<400> 7

Met Ala Ser Ser Val Lys Leu Ala Thr Ala Leu Lys Gln Arg Ala Ile  
 1 5 10 15  
 Leu Thr Lys Glu Leu Ser Glu Leu Asp Asp Lys Ile Gln Ser Ser Leu  
 20 25 30  
 Ile Ser Gln Val Gly Met Lys Lys Ile Asn Asp Pro Asp Lys Leu Tyr  
 35 40 45  
 Leu Asp Tyr Val Ala Lys Ser Gln Glu Leu Ala Lys Leu Val Ser Ser  
 50 55 60  
 Ile Asn Tyr Thr Asn Asn Ile Thr Pro Ile Glu Leu Asp Leu Thr Met  
 65 70 75 80  
 Gly Lys Tyr Asp Asn Thr Ile Lys Thr Ile Asn Asp Ala Leu Ile Cys  
 85 90 95  
 Arg Asp Arg Ile Phe Lys Lys Leu Gln Phe Val Lys Lys Ile Ser Thr  
 100 105 110

Ala Gly Lys Glu Gln Pro Leu Asp Ser Lys Asp Glu Ile Lys Phe Val  
 115 120 125

Ser Phe Ile Asp Val Asp Lys Tyr Asp Thr Leu Ala Gln Glu Leu Asn  
 130 135 140

Thr Gln Phe Glu Asn Leu Asn Leu Lys Leu Gln Glu Ile Asn Trp Gln  
 145 150 155 160

Val Asp Leu Val Glu Ile  
 165

<210> 8

<211> 161

<212> PRT

<213> Candida albicans

<400> 8

Met Lys Leu Ala Glu Ala Leu Asn Leu Lys Lys Asn Leu Glu Arg Asp  
 1 5 10 15

Ala Gly Glu Leu Lys Ser Leu Ile Leu Lys Cys Cys Gln Ala Gln Thr  
 20 25 30

Gly Glu Asn Pro Pro Phe Asp Pro Asn Glu Leu Phe Glu Gln Tyr Glu  
 35 40 45

Glu Ile Asp Lys Leu Ile Thr Asp Ile Thr Ile Lys Ile Gln Arg Thr  
 50 55 60

Asn Asn Glu Ile Lys Phe Ala Tyr Asp Asn Asp Lys Ser Asn Glu  
 65 70 75 80

Glu Pro Leu Arg Ser Met Thr Gln Ala Ile Ala Asp Ile Asp Asp Leu  
 85 90 95

Glu Arg Gln Ile Asn Val Thr Asp Asp Ile Ile His Asn Gly Ile Ile  
 100 105 110

Thr Lys Ser Tyr Ser Thr Lys Lys Ile Ala Asp Val Ser His Val Asp  
 115 120 125

Val Val Ala Tyr Asp Lys Thr Arg Lys Lys Met Asn Glu Arg Leu Asp  
 130 135 140

Lys Leu Lys Leu Arg Ile Gln Ser Ala Asn Trp Glu Phe Asp Leu Ile  
 145 150 155 160

Asp

<210> 9

<211> 6619

<212> DNA

<213> Candida albicans

<400> 9

gttgataaga tctctaataa tgatttgtaa tttgagcgaa tttttatctc ttgttggtt 60  
 tttgtggatg ttgcacataa agctgcaagg acatcaccaa caacaagtag caagtgtggc 120

tagagttaca	aatccgtgta	tggtagcaca	actgatgaca	tttgaataga	tgtcatacaa	180
caaaatatgg	aatagttttg	gataataaac	agcacgtgac	tattgttaac	cagatggctg	240
ttgagaagac	actaagacag	tacaacagat	atctacaaac	acctataggt	aaatgaggac	300
tgccattttc	cttgaaacca	ttttctatta	cttattttaca	ttagttgtat	cttttcatta	360
attcaatttc	attcataaat	atcaaatacc	tagtatctaa	ctacatattg	cctactttaa	420
atgaaaataa	accttggggac	ataataaatt	atatctgata	atcttagtac	tggcccccatt	480
tccataaaga	ttcacctgga	ttaaccaatt	cattatccaa	tttactggag	agtaaaccctt	540
tatctaaacc	ttgagcagca	ataacttgta	aataccataa	gaatttatca	ccatcaaaaac	600
caactttcaa	agattcagta	aactctttct	tggegtcttc	tgagattgt	tctgaaacaa	660
ttttcaattc	acatttatga	tctttagagt	atctatcaat	tacatcttta	acagtaattc	720
tatcagaacc	aatttctaag	gtatctggta	attctcttgg	attaggatta	gtagcaatag	780
aaagtacagc	tttaccaaca	tcttctaatt	tagaaatata	aaattgttga	ttaatgtcac	840
caatgagttt	aacagttctg	tcttcagcat	taataccagt	tgaaccaacc	cattcataaa	900
gaaaagctcc	tggaacagca	aataacgaag	ttataatata	aacaacttta	actcctgatt	960
ttctgacatt	ttctgaatgt	tgagttttta	ttcctaaaaa	ccctggagca	tattcatcaa	1020
cttttaggaat	atcagtacca	aattgtgatg	ggataaataa	ttttggttta	attgcatcaa	1080
ctaaattttc	gatattggcg	aaagcctctg	gattttggacc	aattaattca	ataataacat	1140
cagtagcaga	taattttttg	ctcaaaagttg	atttaattga	ttcttcattg	atttcagaaa	1200
caacatatct	aattttgtca	ttcttagttt	ctgggttcctt	tctggtaatt	gccttgattg	1260
ggaagttgat	cttatcatca	aaaataccag	aattgatagc	ttcaagaact	ggtttaccta	1320
agaaaccatt	caaaccgatg	atagtaattg	agactttaga	cattgtgata	gatagattat	1380
atagattaat	tattagataa	gcttgtgtaa	ttgatcaatt	gcttgattaa	tgagattgga	1440
aaacaaaaaa	ttacaagcca	tgttgaatgg	aggaaacacg	tctatttata	atgggtttga	1500
ttcaatgtga	tgcttaatat	gggagtgggg	ggttatgcaa	tgtaaggaga	gacgacaaaa	1560
catacttagc	taaaaacaca	aacacacatt	gttgccatag	ttaaatgtgg	aattaaatgg	1620
aacaatcttt	tcccgtaaaa	tgtaaaagaa	ggaggaaaaa	catacaccaa	gaaattgtgg	1680
cgtaatctga	aattctttgt	ttctctcttt	ctctgtttta	tttgtaatca	aatatttttc	1740
tcattacata	atatgcaagt	gatgattaat	aatcaatatt	tgtttatcag	ttatatctat	1800
ttaatccttg	tatttataat	ttcataacaa	atcaataaca	acaccgcta	cagccacatc	1860
acaatcaatt	tactggtaac	ttattttgta	tctacatatt	acctaagatt	gtacagaaat	1920
tgttttctgt	tactaaattg	tattggtaat	aatcttacta	tgagtaaat	agtgttgcta	1980
ttataattgt	ggctagtgtg	tatactgata	tcaatttaact	cgtattaata	atatattagg	2040
tgtagcaagc	ttgatattct	tgacacagct	gttattttgt	tacgccacat	tagattttct	2100
aaccacaact	tgtaaggtag	tatagacaac	taaaacccta	tacagggctc	tcttaatcct	2160
tcttccaatg	tctgactact	aagtaccatc	ataccaccag	ggtaggttag	atatatagac	2220
tctcttggat	gcccttgctg	tctactccta	aagaggttta	catactcaac	tgaaaaatac	2280
aagctattaa	cgtagttaa	gtattacaca	actttgtata	gctgactttg	caaccctcag	2340
attcggtgaa	ttttttttgt	aagtaacaat	ctgtgtgtct	tggtaaaatg	ctcaagcctc	2400
agttattact	aatcttatta	cttatactgc	caattctaaa	acatcagca	aattcattat	2460
taagcattgg	tttttagatt	catttgggta	gtaaaagtgc	agatacgata	tacataacta	2520
agcatagact	ataataatcc	gaaactaaaa	gttgcaatat	tccgatatcc	ggtatccttt	2580
ccgttgaatc	ttttgtattc	taaatattaa	tacaaggtat	agcttgatta	attgttggtt	2640
tatgtagatg	acagtacttt	acttattggt	tgtaaaataa	ctaaagaagt	acagaattta	2700
tttcaaaatt	ggtatccaaa	ctcaacccaa	ataataaatc	aaccacccaa	attgatatta	2760
tctaaaacta	caaataaaga	tatacaaatg	gtttagccta	ctgaaagtat	acataatagc	2820
atcagttaac	tctctgtagt	tcaaccaact	aaaaataacc	tttataatgg	ttctcatctc	2880
tagttgtatc	attctgttcc	tcactctcgt	catcttcata	ttcgtttgaa	ttatcttcaa	2940
cagtatataa	cccaattctt	cttcgattat	tattattggt	atttgatgat	agctgtttat	3000
tctttttaaa	tttcttgaaa	ttttctcttc	cagcatatct	cgattcagta	tcacatcaa	3060
cactcttatt	tttgttacgt	cgatttctag	gattaaacta	atcaacaatt	tccacaattg	3120
ccaaattttg	taaattttca	ttaatttcaa	attcagattc	caaactctaa	tttgatgatt	3180
tattcaattc	ttgaactttt	aaaactgcat	ctactaatga	tacttttggg	gtaaaattttg	3240
gaattttaaa	tggtttttga	tcacatttag	tagtttcatc	tttattctta	cgttttgatc	3300
cgattgcagg	aatttcttcc	tcgtcttctc	gttcttcttc	ttcattttta	attgatggaa	3360
aattggaaat	agacttttgt	gaattggaa	ctgggtttaga	tttcttctgt	aattgatcag	3420
ttgatgatga	taaattcgaa	tcacccaaac	tattaaagtgc	atcttcttgg	ttgattttac	3480
gactttgcgt	atgagaattc	aacacttttt	gattttgctt	tatcgttgaa	atttttttaa	3540
tgggatcact	taattgtggt	atttcttgag	ttattgaatc	attatcacca	ctggatttaa	3600
ttttctgact	atttaattct	ggtgatgtgc	ttttattgtc	agtggtccaa	attgatactt	3660
taggggatgg	atcacgtgta	gatggtactg	gtgtagttgt	agttgttggg	gttgctgtac	3720
ctaaggagaa	gaaatgtaat	ctatcaactt	tttcatattt	ccttctcttt	ctccctgaag	3780
ggtgccgcca	cctcaggtgt	ttgtgttgta	actgatttat	catctgatgc	taacatagtt	3840
gttgaggtac	gtttcaatta	aatcaaaata	tttttgggaa	gatctagtat	tgagccttct	3900



```

ttaattgatt gccataattg atcttctgtt agagtttggga aattatgtct attcaaagtt 3960
gttgccattg tagttggaac aattaaaatg aattgattat cacctatttt cgatttcaat 4020
acgtcttctt gttgttgat taacacttta aaatctataa ttattgtttc caaccaattt 4080
attctctcca atgatactaa tttaacatta tttaatggct ttaatctata acttttcggg 4140
aaatagtgat catccttttt tggttaagat ttcttataat ctatatccaa aagccaatcc 4200
ttaacactat tatacccatc ctcatatca tgacctataa ctttttgtaa ccattgtgaa 4260
tttaaaattg ggataccctt aataattgcc aatttgaatt catctttttg ttcttcctcc 4320
ttgatatcgt tattatcact atcataataa taatgagtgg catccattat attatcaacg 4380
atttgaatat caataccatg atcttttatt attgttttgt aatgattttg atcacatttt 4440
tcattcgcac taatatataa tttaaaggga atccattcaa tgaacaatgg agtatcttga 4500
gttattattt gtcctgtttg tttatcaatt tgtggaatag tttcaatttt caatttgaat 4560
tcatttgtga aatcaatttc ttctgtttta ctatcttttg ataataattt atatgttttc 4620
ccatttataa tcgttttcgc tcgactacat attctaataa ttaataatgt tttatcacta 4680
cctattgaag aagaagacga tgacgatgat ggtggttggg tattgcctat gttgggtcca 4740
gttatttaatt caatcaattc tcttgatact ttagggtgacc agaattttat atcagtattt 4800
tcatcacgtc ccacggtata atttgtttct ggtgataaat gtttatattc taatgaattg 4860
ttagtgggcc tgggggttaa taaatatcaa aaatcttcag tttctaact acctcctctt 4920
ggatcatctt tatgatattg tatccacata gtttaagtga catttgcttg tcaatgaagt 4980
cttcctaagc gaatttgaac tgaaaaaaaa cctatcgcgga tccctttaca actttaacaa 5040
acgtataata aacaggttgc ataaaaacct ctacaaacca ttcagaatct tactttcata 5100
aatagtattt tattatgaat cgtttcttta aactagtaat acttgatca ataataatc 5160
cctaagcttt tatgttgtaa ctggattaat atcagagtgt taggcgtggt cacgtgacat 5220
agatagaata agagtgcgaag ggaacaacat taattagact tgatacttat tgtattaaga 5280
taaagtgaat tttacaataa caaatggtga atatagttaa tcagttcata ctagacgatt 5340
agttgaaatt ttttatgtt tgtttggata atcgtttgtt ataaaaatag aaatcgtcac 5400
tttttttttg ctcaattgtc ctgttcattc cattcaaggc tgtaactgta aatgttaaac 5460
taaatttcat ttcatttcat tcaacaaaaa aatatcatat tctattttaac ttcaacaaac 5520
taatctcaag aaccagctta ctctctttta taatactaca acaacatata ttatacaact 5580
aatcaaaagt ggggtattaag aaaatgtttc aaaagaaaga accaaccgaa caagaaattc 5640
gtaaagaatt aagtcgagtt ggcatttcta caagatcaaa taatactcga caagagaagt 5700
ttggtgcatt taaaaattat gctcaagaac gagctaatat gaaaccacaa ttaggaccag 5760
ttggtggtaa tccttatgcc aatattaatc ctgggacca caataataat aataatccat 5820
atgccaatga taatgggaat aatagtactg gcaaccccaa caacaacaac ggtggtaacc 5880
cctatggtgg tgggtgtact aataataatc cttatggagg ctctggtggt aatggaagag 5940
gatcatcacc tagtccttat gcaccgacta catcaacaac tactagatca tctaatecat 6000
atggaacaaa taatggttct agatcaagtc aaaacacttc tagtccttat gccaaatcaa 6060
ctaacaattc atcatattct aactcaccgt attctggatc aactgtaaat aatggtaatc 6120
gtggcgcca tagcaacaac agcaatagtt ctgctggtgg taaccttat gctgccggtg 6180
gtagaagttc acaatctcaa aattcacgag acaatgtata tacagctcct gccactcgta 6240
catcaactag acaaaactcaa ggatattggag gtggtgatac cgattcgact cttgacctta 6300
atgccattcc atcacatcaa atgtttgata ataagaaacc gatcaaaaaga aatcaacaaa 6360
gttcacaaca acctgccaat gattataatt tagattttaa tgatgaatat ggcgaagaag 6420
aagacttgaa ttggatata agtgaagtac ctgaagaaca acaacaatc aattctgaag 6480
atgaagaagt agaagcatt aaacaagata ttaaatttgt caaacaagaa tcagttcaaa 6540
gtaccagaaa tactcttaga atggcacaag aagctgatgc atcggtact aatactttag 6600
gaatgttagg atcgcaact 6619

```

&lt;210&gt; 10

&lt;211&gt; 7965

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 10

```

tagctaagta tgctttgtcg tctctcctta cactgcttaa ccctccaccc ctctatttgg 60
cttcacattg aatcaaaacca ttataaatag acgtgtttcc tccattccac atgacttgta 120
attctttgtt ctccaatctc atcaatcaat tacacaagct tatctaataa ttaatctata 180
atatctatct atcacaatgt cttaaagtctc aattactatc atcggtttga atggtttctt 240
aggtaaacca gttcttgaag ctatcaattc tggtattttt gacgataaaa tcaatttccc 300
aattaaagct attacaagaa aagaaccgga aactaaaaat gacaaaattg aatacgttgt 360
ttctgaaatc aatgaagaat caattaaatc aaccttgagc caaaaattat ccggtactga 420
tgttattatt gaattaattg gtccaaatcc agaggcttcc gctaataatc aaaaattaat 480
tgatgcaatt aaaccaaact tattcattcc atcacaaatt ggtactgata ttctaaagt 540

```

tgcagaaat gctccagggt ttttaggaat caaaactcaa cattcagaaa atgtcagaaa 600  
 attaggaggt aaagtgtgtg atattataac ttcgttattt gctgttccag gagcttttct 660  
 ttatgaatgg gttggttcaa ctggtatcaa tgctgatgac aaaactgtta aacttatttg 720  
 tgacattaat caacaatttg atatttctaa attagaagat gttggtaaaag ctgtactttc 780  
 tattgtact aatcctaate caagagaatt accagatacc attagaattg gttctgatag 840  
 aattactgtc aaagatgtca ttgatagata ttctaaagat cataatgttg aattgaaagt 900  
 tgtttctgaa caatctgcag aagatgccaa gaaagagttt actgaatctt tgaaagctgg 960  
 ttttgatggg gagaaattct tatggtattt acaagttatt gctgctcaag gtttagataa 1020  
 aggtttactc tccagtaaat tggacaacga attggtcaac ccagggtgagt ctttatggaa 1080  
 atggggcaag tactaagatt atcagatata atttattatg tcccaagggt tattttcatt 1140  
 taaagtaggc aatatgtagt tagatactaa gtatttggtt ttaataaatc agattaaatc 1200  
 aatgaaagga ttacaattaa tgtagatgag gattagaaat tggtttctaa tagagaactg 1260  
 ttgtactctc ttggtgcctt ctcaacagcc atctggttaa cagtagtcac gtgctgttga 1320  
 tcatccaaaa ctattccaca ttttgcttta tgaaaacctt ccctaattt gccaattgtg 1380  
 ctaccataca cggacttgta attccagcca cactttctta cttgttattg ttgatttctt 1440  
 tgcagcttta tttgcaacat ccacaaaaac ccaacaagac aaaaataacc tcaaattaca 1500  
 aatcatcatt aaagatctta tcaacaacat gtccctttt tgataattaa caccttaaat 1560  
 taacctcttg tggggttgat caaaccttga atttgaacct taagtgaag taagaacgaa 1620  
 gttaaagcat ccaagtcttg tgaaaagca acatgtaacg gcattttagt ttcgaaccga 1680  
 gttaaaaatt ggaagtaaga ataagagaaa ctgctgacg gctgtgcggc atattattta 1740  
 aggcaaatg acgatggcaa aaaaaatctt taatttttcc gtcttggtta ttaaatatcg 1800  
 cgaaggcatt cctttaaagt aatgccttta tgaattgtca aaatgcccta ttttaaaatg 1860  
 ccttattact ttctattgag cttatcatga aatttggtta ttactggcac ggcaacaatt 1920  
 taatcagagc atctgaaatg tgggtgatta attgcagagg gcggttagat ataattttag 1980  
 atggattgat ttgcctaatt aggaatttgt aactttataa acctctttat aatttttcat 2040  
 cttttcattt ttttcaaaat aaaatgggtt tttcacagaa gtaaagatat tgttcatgct 2100  
 tattgtgcta acctgatgtt aatctcttta ttttaatttg acttcttctt caaacctgtg 2160  
 aacacggcac caatgaatga taaaattgat gatatttagt taccagcaag caatgatgac 2220  
 agtattccca agaagaagag acaatggcta cagcaactgg agtaaaacaa gtattttagc 2280  
 tacagatcta acctttttgt aacaccccaa ccattctctc tttttttgct atattgaaac 2340  
 actgattttt cttaaattat ttacccaact cctaattgaa tataacaacg aggtgcttct 2400  
 acttctccac atttccaaca ataaaacatg aatcaacatc aatattgttg caattaatca 2460  
 aactcttaac atctccacct cccactaat gatcgatatt atatcaaac cattggaaat 2520  
 ttagtttggg tttcatttga atttcggtca agaaaattaa aagtaaaaaa gaaaaaaaaa 2580  
 atttattatt attattcggg tcgatattgc cgcaaaaacca aattccatca tcatttcaca 2640  
 atataatata aaaagcttct aatcttacac ctgcaaaaaa gtttcaattt tttttataa 2700  
 aatatttatc tatattttaa ttgttacatt tattctttac ttctaataca aacaactata 2760  
 tatcaatatt atgtttgaag taggtgaaaa atatcctgtt gaaagcagca gtatttcaaa 2820  
 tgacatagaa tctcgtggtg ttcaacctat aacatccctc aaagacaata aatcaatagg 2880  
 aatgatagag aaagataatg atgatctatc atgtgaacaa tatagtactt gtgatgaagt 2940  
 caaaagagat ttaaaagcaa gacatgttct tatgattgcc attggtggtt caatagggtac 3000  
 agggttatct atatccactg gttctttact tcacaccact ggtccagtaa tgtcattaat 3060  
 atcattttta tttgtcaca ctttagcata ttcagttaca caatcacttg gggaaatgac 3120  
 aacatatatc cccgtttctg gatcatttgc ccaatttata actcgttggg tttcaaaaag 3180

ttgtggtgct gctaattggtt gggttatattg gttttcatgg gccataacat tcgctttaga 3240  
 attatcagtt gttggtcaag tcatacaata ttggactgat gctgtaccat tagctggttg 3300  
 gatttccatt ttttctgtct tattaactac atttaattta tccccggtga aatattatgg 3360  
 agaggttgaa ttttggattg cttcaactaa agtaattgct attgttgggt ggctcatata 3420  
 tgcattttgt atggtttgtg gggctggtta aactggacca gttgggttcc gttattggcg 3480  
 gaatggatat gcatgggggt atgggatgat agttctgaat aatggtaaat acgccatttc 3540  
 tttcattaat ggtcttatca atgctgtttt tactttccaa ggtactgaat tgggtgctgt 3600  
 tactgcgggt gaagcttctc caagagcaat ccgtagtgca attaaaaaag tcatgttcag 3660  
 aattttggta ttctatgtct tgtgtatgct tttcattggt cttttgggtc cttacaacga 3720  
 tccaaagctt actcaagatg gtggttttac aagaaactct ccattcctta ttgctatgga 3780  
 aaattccggg actaaagttt taccacacat tttcaatgca gtgatcggtt caacaattat 3840  
 ttcagctggt aattccaatg tttattcggg atcacgtatt ctttacgggt tagcccaagc 3900  
 tgggtgagct cctaaatttt tccttaaaac taacaaagggt ggtgttccat attttgctgt 3960  
 ctgttccact gcggcatttg atgcattggg atatttagca tgttcggaag atggtaataa 4020  
 agctttcact tggttattga atattattgc cactgctgga ttgatcgctt ggggattcat 4080  
 ttctgtgagt catgtcagat tcatgaatgt tcttagaaaa agaggtttaa gtcgagacat 4140  
 tttaccttat aaagcttttt tcatgccata tagtgcata tatgccatta ttattatatt 4200  
 cattgttgtg ttgattcaag gtttcacagt gttttgggac ttcaatgcta gtgatttctt 4260

cactgcctat	atatccgtga	tattatttgt	tgttcttttg	attgggttcc	actttttctt	4320
ttacgggttt	ggtaaagatt	cttttaaatg	ggaaaacata	ttaatcccat	tggatgattg	4380
tgatattgat	tctgggtgta	gagatattaa	tgatgctgaa	tttgatgtac	ctgaacctaa	4440
aaatgtttgg	gaaagattct	ggttacttat	tgcttaatct	taatttatat	aatttaatac	4500
ttaatggaca	tagagctttt	cagcgattat	aatagaaact	gattataact	atttaattta	4560
aatctattta	caaattttct	tgaagagtgg	ggcgccgggt	atacaaaact	aacaacgtaa	4620
taattctgtt	taccaactct	aaccaacaac	aattttaccat	caatcaaagt	attatcaaca	4680
tcaaataata	taacatcatc	tggatcctca	atttgatttc	tgtccaatcc	catgtataca	4740
ccaccagctt	taatcaatct	tctcatttca	ccctttgatt	taccaacaat	atcagctaata	4800
aatgcactca	atttgatttc	ttcatcaggt	gaaggcttat	ttctttttaa	caaaataacct	4860
gatcttttaa	aattttctat	caatttatca	gcgcttacat	tatcattaaa	tggttgatct	4920
ggagtaggga	ataaaaatcc	cgtaataaac	gccatttcgt	caccaacacc	aacaccatgg	4980
atcaaataca	caacttcacg	tgctaaaaca	cgttgagcaa	tacgtaaacc	aggatcactg	5040
ttatgtttag	gtaataattc	acottcaatt	acattcaagg	gcaataatgt	gaacactttt	5100
aataatttgc	ccactatata	atctggaaca	ttaatgaaat	attgatacat	ttgataagga	5160
gtgggtcaaac	tgaatacaat	aaatactgca	ttccagcggt	atttaccaaa	tttctcacca	5220
ctagaagtag	tcaataatgg	aacagtaagt	ccataagcgt	catgtttctt	gccatgaaat	5280
ttcttcaaac	gtgaaattaa	atcaatacca	gcagtaatat	tccccattg	atcatttcct	5340
ccaacttgca	tattaacatt	ttcatccttg	tataaatgcc	aaaaatcata	agcttgtaga	5400
atctgatagg	taaattcatt	gaatccaatt	ccaccaggtt	ctaactctga	ttgaatggaa	5460
tcacgtgcta	acatcgaaact	aactctaata	tgtctaccat	atgtagccaa	aaattccaac	5520
atcttcacgt	tttcccacca	tgaagcatta	tttaccgatg	tagtatcacc	gactttttca	5580
gtcatgggga	attgtcttga	tttggcatat	tctatcccat	tactcaaaaa	tgtagaaatt	5640
tgtcgttggg	tttctgtcac	attatcttca	acttcaactt	catcaactct	gttacgctct	5700
gttttctctc	cacttggatc	tccaacaagt	ccagtagcac	ctccaacaag	tccaacaaca	5760
tcattaccac	tcattttgaa	atgtaataac	accattaatg	gtaataaatt	acctaataatg	5820
agtgatgatg	cögtaggatc	agcaccacaa	tataatttga	atttgtgatt	agaaccacgt	5880
ttagtcaatt	tatataaatt	atcatcggtt	attgattcaa	ttaaatgtcg	actttgtaaa	5940
tatttctaata	atgaattgtc	gggattgggt	tctgggtgta	aatctttggc	ttcagtcaat	6000
tcatatagtg	taggaatgat	agtgactgga	tctcttgcaa	ttgttgaatt	aaatctagca	6060
agccttctaa	tcaaaggtat	gtttctggtg	tgtgttttca	acataattact	tgatgtctgg	6120
ttgaacttct	ggttgtcgtt	tcttcgattg	aattttttct	tgtagcttca	ttagcgggct	6180
tatttgctat	tcgcggttta	atttttaaag	aaagccgcaa	attcaaatcc	aaatccatct	6240
caagctgaga	tttttcttta	attttttttt	ttttcacttt	actgatataca	ttctaatacat	6300
taaacataca	aagctcctaa	accaatgaca	gätcagatta	aaggggatga	acataaccaaa	6360
gtaaatggta	gtaaaaaaag	aaagagaaaag	agaaataaga	acaagaaaaa	tgacaacaac	6420
acaccagtag	agacttccga	accaataaccg	acäcctgttt	atgaagatga	tatccatcga	6480
caaaataaaa	agttcaaatt	caatgaggaa	ggggaaatgg	agaaacctca	agagtcacct	6540
gaagaggaaac	aactgaattt	agtggcagat	caaggggaac	cttgactga	agaaccttta	6600
ccacaacatg	aaggtttcga	agagattgaa	gtaactgacg	acatagatga	aacagaagaa	6660
ccagagaatc	ttccaacaag	gactcaacaa	gaaaaacatc	aacacggtaa	gaataaattt	6720
aaacaaaagc	ttgaattcaa	aagaaaaacc	gttgatatata	aagatcaaga	tgatgaagac	6780
gatgagggaag	aaaataatac	tttcaatttt	tcacaaaatt	cgtttcaact	tgcaaccacc	6840
gcccacaat	tgttacaaat	tagagagaaa	ttgcccattt	atcatcataa	ggataaaatc	6900
attgaatgca	ttaataataa	tcaagtcact	atcgtcattg	gtgaaaccgg	ttcaggtaaa	6960
tcaacacaaa	tccctcaatt	tttaatgcc	gaaaacccaa	aaatgattgg	cgtgacacaa	7020
ccaagaagag	ttgccgctgc	ttcttttagca	gcaagagtaa	gtgaagaata	tggatgtaaa	7080
ttaggtcaag	atgttgggta	tcaagttaga	ttcactaata	tgactaacag	acaaacaaaa	7140
ttgaaatatt	taactgatgg	tatgcttcta	cgagaaatca	tgcttgatct	gaatttgact	7200
aaatattcaa	caattatcct	agatgaagcg	catgaaagaa	ctattttgac	tgatttaatc	7260
atggggtttt	tgaacaaaat	tattacttct	ggtaaaagaa	aagatttgaa	aatcgtagtt	7320
atgagtgtcta	ctttgaatgc	cgaattattt	agtaatttct	ttgataatgc	tcctatttta	7380
tacattgaag	gtaaaatgta	tccagtttca	caattctact	tagatgtga	atctgaagat	7440
attgtggata	ccatgatcag	aagtataatt	caaatcaatc	ttaatgaacc	cgaaggggat	7500
attctttgct	tcttacctgg	gcaagaggaa	attgataatt	gtgttaaaaag	tttagaaca	7560
ttagcacctc	aactacctag	ggaggcacca	ttgattgttc	ctttaccttt	atatgcagct	7620
ttatcacctg	gccaacaatc	taaaatattc	gaaaaattac	ccaaggggaag	aagaaaagtg	7680
attttgcgga	caaataattgc	tgaacatcc	attactgttt	ctgggtgttaa	atatgttata	7740
gattccggat	taaggaaaat	taaagtttgg	aaacataatt	taggactttc	tacattattg	7800
actaccctta	tttcacaagc	ttcagcaaga	caaagagccg	ggagagcagg	tagagaatct	7860
gaaggtaaa	tattcagatt	atatcctgaa	tctacttata	tggcacttcc	aaaacaacaa	7920
gaatctgaaa	ttaaaagaaa	tgatattatt	ttaccagttt	ttgac		7965

<210> 11  
<211> 5158  
<212> DNA  
<213> Candida albicans

<400> 11

aataattatc	attagtcaat	tcaacaacta	taggagattt	agcagcaatc	tcttgagagt	60
tccttctcgc	acatctaaca	acaacttgga	tatttgacat	tgatgataat	gctgctatga	120
gtattaattt	aactgaaatc	acaagatgaa	gaatgaaaac	aacaacaaca	aagagaaaga	180
gtttggcaac	gggagaggaa	gagaaagtgt	aaacaaaaac	aaacaacccat	aaaaatttac	240
accataaaaa	aaaattagaa	gtcgtgattg	aactatatgc	aggccactat	aagaagatat	300
taaaactact	ctgattgaat	gaatgaatga	ttatataaat	ccctcttttc	tctcaactta	360
tagccttaat	caaagaaatc	atcgtcatct	tcacatcat	cagcagctat	atttttcccta	420
ttgaaagtta	tttttgtttg	ttgcttttgt	tgactatctc	tacttaatcc	atctaattatt	480
tgacaatat	cttttccacc	aagtttttga	tgtatttgac	ccattgaata	taatttaata	540
atataatttt	ctactctgtg	agctctatcg	ggcttaacaa	tctttacacg	acttaattctt	600
tctctagctt	cattagttaa	gactcgattt	aatatggtaa	tggtcatatt	ctcttggtgcc	660
agatcttggtg	cgccaccoga	agaagaagaa	gatggattgg	tactactgcc	acctccggca	720
gcatttcttt	gtaattctgc	taatcttgct	tgtcttatag	catttaattc	tgctcatcc	780
ataatgtata	gttggaatg	aatggagaag	gagtagtatt	aaaataatta	gtgtggatag	840
agtagtcagt	cagtcaacca	agtgaatagg	gaagtaagga	aaaatttttg	tcacatttaa	900
cacgaacttc	ttgatcaaa	aagaagaaga	agaaattttt	tttctgtcat	cacgtgcacg	960
acctttaaat	caattgacaa	ttcaaaaatt	ttgaacaaca	acacaacaca	actcattctc	1020
tctttctctt	tctctctctc	tctctgtgaa	aaaaaaaaaa	aaagtaaagg	acaataaaga	1080
aatcaaacaa	tcaattaaac	aaagttaaaa	caggaaactt	ttctattcaa	gttcaaattg	1140
aaagagaag	agaaatagaa	agaaaaaaa	aatttagttc	aaattggaat	cttgctttat	1200
ttagtttcat	ttctatatat	cttgctctca	tatactatca	acatttagat	tgatttgaat	1260
ccagaatcaa	caatttcaac	aattcttcag	attttgatat	agtgatttct	atttgacata	1320
ctttactact	accaatacag	tcacataatt	acatatataa	atatattaag	agtgggtttt	1380
cggaaacttt	tcttcttaga	tttaatatag	aatccctttc	ccctaatttt	ttttgcatca	1440
acattttact	aaaaaactta	accccaccaa	ctcctaacc	taatatttcc	cctttctttt	1500
tttgcatata	agactccac	aatgagttca	gataaatcaa	atttactaaa	aaaatacaag	1560
attgtctttc	ttggtgatca	aagtgttgt	aaaacatcat	taatcaccag	atttatgtat	1620
gatacatttg	atgaaactta	tgctgccacg	attggaattg	attttttatc	gaaaacaatg	1680
tattttagaag	aaggtaaaac	cattagatta	caattatggg	atactgccgg	acaagaaaga	1740
tttcgatcat	taataccttc	atataataga	gattctcatg	ttgcagtaat	atggttatgat	1800
taaaccaata	aaaaatcatt	tgataatctt	gataaatgga	ttaaagatgt	taaattagaa	1860
cgaagtgatg	atgtaataat	agattatgct	ggtaataaac	tggtatttagc	tagtgataaa	1920
cgacaagtta	gtttagatga	tgttgaaaat	ttacaaatta	aaattgggtgc	taaatttttc	1980
attgaaactt	caactaaagc	aatcataat	gttaatttat	tatttaaaaa	aattgctcaa	2040
tcattacctg	attttaatca	agattccaat	gataaatcaa	atgataataa	taataataat	2100
aataataatc	aactggaaac	tattgatata	actattgata	atactgcacc	aaatcctcaa	2160
ggtaccagca	catgttggtta	gactagaatc	ttagtgtaa	aactaataaa	aaaacagagc	2220
aatgggtaga	taataattcta	agtatattaa	cagttcatat	aacaacaacc	cacacacaca	2280
catatatata	catatatat	atatcactta	tttaatacat	tagatcaaat	tccaatttg	2340
ccgactgtat	acgaagtttt	aatttatcta	atctctcatt	cattttcttt	cttgctttgt	2400
catatgcaac	cacgtcaaca	tgtgacacat	cagcaatctt	cttggtcgaa	tacagttttg	2460
taataatacc	attatgaatt	atatcatctg	tcacattgat	ttgtctttct	aaatcatcaa	2520
tatcagcaat	agcttggtgc	attgatcgaa	gtggttcttc	attagactta	ttatcattat	2580
cataggcaaa	ctttattttca	ttattgggtc	gttgattttt	aatagttata	tcagtaatta	2640
atttatcaat	ttcttcatat	tgttcaaata	attcatttag	atcaaattgga	gggttttcgc	2700
cagtttgagc	ttgacaacat	ttaagaatta	atgatttaag	ttcaccagca	tctctttcca	2760
agttcttttt	taattttaat	gcttcagcta	atttcatctt	ctctcacaaa	atataatata	2820
actcaattat	tggttgtaag	atttatatga	ataaagtata	tgaaaatgaa	aaaaaaatgg	2880
ggtgagaggt	aaatgtatcc	gaatttataa	ttctgttgat	agcggagaaa	agtataattt	2940
tatttttttt	ttggtagttc	ggttgtagtt	cctctttgct	ttattcccct	atgcaccctg	3000
gcatacacaa	aagtcaattg	attgctttct	cttgtaagg	cttttggggt	tggggttgga	3060
gtaattgtcg	ttgttgttgt	tgctgttctt	gatacagtg	aatagagata	tgactaattg	3120
gtattggtat	gtttatgttt	acacaacaca	tatgagtc	cgaaaaatca	attggcttga	3180
tctgacttct	cctggaaacta	aattctataa	tttcatcaac	aattgtaggc	aaatgtagac	3240
aaatgtgtgt	gtttcgtcta	gctcaatata	accatcaagg	tttggttaagc	ctccttcctt	3300
tattattttt	gcctcttgaa	aggcattttt	gatgtaacaa	agtgattcta	caattgttgc	3360

```

gagcaaatta ttggcaaaca tcttttgtga aagaatcata accttccatt cgtttgttcg 3420
ttttgttagc tcattggctg atgggttctt tagttgctat gaatactgct gctctgtttt 3480
caaaatcctt ttgttgggaa ggttctaccg attgaggttt aacttgatt atcgtgtaag 3540
tgtgttcctg actccgaatt tttgtctata aatagacctt gaaaagttca ctttttttca 3600
aatttttttt tattcccttt tctttttttc taatcctcat taacaaatca tattcaaaca 3660
aatcaatcat tttatgcatt gagtcgtatt aattgttgtt tgttggttat agcttgttgg 3720
ttgattgatt ggttgggttg tagtataaac attttcatta ctctaattggc ctccctcagt 3780
aagttggcta cggcacttaa acaacgtgct atattgacaa aagaattgtc tgaattagat 3840
gataaaaatac aatcttcatt gattctgcaa gttggtatga aaaaaatcaa tgatccagat 3900
aaattgtatt tagattatgt tgctaaatct caagaattgg cttaaattgg atcatcaata 3960
aattatacta ataataaac tccaattgaa cttgatttga caatgggaaa gtatgataat 4020
actataaaaa caattaatga tgcattaatt tgtcgagacc gaatatttaa aaaattacaa 4080
tttgtgaaaa aaatatcaac agcaggtaaa gaacaaccat tagattccaa agatgaaatt 4140
aaatttgtat catttattga tgttgataaa tatgatactt tggcccaaga attaaatact 4200
caatttgaga atttgaattt gaaattacaa gaaataaatt ggcaagtga tcttgttgag 4260
atataaaaag gatagtgggt ctggatcgcc attgataata tcttttactt gttactttat 4320
gtaaaaggat ttaaaaaata ttgttggtac tactcgtttc ctccctccca aatcgaataa 4380
tagaactata gaaccatata cccctataaa ttattttatc tgattttatt agttataaag 4440
tacaaatcta ttatcaattg ttttattatt tagtattttc ctccaaagtt ttgaactttt 4500
gttttttatg gttctagttc tttattcttg tttttgggga tttagggttg ccgcttgatt 4560
tgttgaactt taattgatgc tttgtttagg catagtaatc aagaaaagga agataatgaa 4620
agggttaggga atgagtagga gggcgggttc ggggacaata tacatgtata gttacgtaca 4680
ttaatgtaaa tatattctta aaattcctag ttgttaaatt aattgatggt gttgttgtct 4740
ttgtattttt aaagtattca aaaattttga gtcaatttcg ttaccaaatc ttaatgaata 4800
gtaacacgtc taaccaaatt tcaacaaaaa gtttcatacg accaacaact tatatgcttt 4860
tcagtatgta tatatcttcc atatttttat ttgtatatga ttgaattgat aattgtaata 4920
gagttaaaag aatgaagaag aagaagaagt gggtttttgc aaccaacaga acagttaggt 4980
tattcttgtg tacacgacca gatcaaatat gtatgtgaga gagagacgga aatagaattt 5040
tcttgaaaga aaaaaaaaaa aaaatttctt tctgtttttt ctctcgcccc gtgtgggttg 5100
gtctctctca ctgttggtga attcgtacca acaattccgg agccaaattt ctttcacc 5158

```

&lt;210&gt; 12

&lt;211&gt; 968

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 12

```

cttctctcac aaaatataat ataactcaat tattggttgt aagattatat agaataaagt 60
atatgaaaat gaaaaaaaaa tgggtgaga ggtaaatgta tccgaattta taattctgtt 120
gatagcggag aaaagtataa ttttattttt tttttggtag ttcggttgta gttcctcttt 180
gctttattcc cctatgcacc ctggcataca caaaagtcaa ttgattgctt tctcttgta 240
aggcttttgg ggttgggggt ggagtaattg tegtgttgtt tgttgctgtt cctgatacag 300
tggaatagag atatgactaa ttggtatttg tatgtttatg ttacacaa acatatgagt 360
caacgaaaaa tcaattggct tgatctgact tctcctggaa ctaaattcta taatttcac 420
aacaattgta ggcaaatgta gacaaatggt gtggttctgt ctagctcaat ataaccatca 480
aggtttgta agcctccttc ctttattatt tttgcctctt gaaaggcatt tttgatgtaa 540
caaagtgatt ctacaattgt tgcgagcaaa ttattggcaa acatcttttg tgaaagaatc 600
ataaccttcc attcgtttgt tegtgttgtt agctcattgg ctgatgggtt ctttagttgc 660
tatgaatact gctgctctgt tttcaaaatc cttttgttgg gaaggttcta ccgattgagg 720
tttaacttgt attatcgtgt aagtgtgttc ctgactccga atttttgtct ataaatagac 780
ctagaaaagt tcactttttt tcaaattttt ttttattccc ttttctttt ttctaactct 840
cattaacaaa tcatattcaa acaaatcaat cattttatgc attgagtcgt attaattgtt 900
gtttgttgtt tatagcttgt tggttgattg attggttggt tggtagtata aacattttca 960
ttactcta

```

&lt;210&gt; 13

&lt;211&gt; 456

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 13

```

atgtccgacg aaagaacttt tattgctatc aaaccagacg gtgttcaaag aggtttaatc 60
tcatctatct tgggtagatt tgaacaaaga ggtttcaaat tagttggtat taaattggtt 120
caaccaactg aatctttatt gagaactcat tatgaagatt tacaatctaa accatttttc 180
ccatctttat tatcttatat gttatccggg ccagtcttag ctactgtttg ggaaggtaaa 240
gatgttggtt aacaaggtag agccattttg ggtgctacta acccattaca atctgctcca 300
ggtaccatca gaggtgattt tgccattgat atgggtagaa acgtttgtca tggttctgat 360
tctgttgaat ctgctaacaa agaaattgac ttgtggttca agaaagaaga attggttgaa 420
tataaaccag ctttgttcgg ttgatctac gaataa 456

```

&lt;210&gt; 14

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Candida albicans

&lt;400&gt; 14

```

Met Ser Asp Glu Arg Thr Phe Ile Ala Ile Lys Pro Asp Gly Val Gln
1 5 10 15

```

```

Arg Gly Leu Ile Ser Ser Ile Leu Gly Arg Phe Glu Gln Arg Gly Phe
20 25 30

```

```

Lys Leu Val Gly Ile Lys Leu Val Gln Pro Thr Glu Ser Leu Leu Arg
35 40 45 50

```

```

Thr His Tyr Glu Asp Leu Gln Ser Lys Pro Phe Phe Pro Ser Leu Leu
50 55 60

```

```

Ser Tyr Met Leu Ser Gly Pro Val Leu Ala Thr Val Trp Glu Gly Lys
65 70 75 80

```

```

Asp Val Val Lys Gln Gly Arg Ala Ile Leu Gly Ala Thr Asn Pro Leu
85 90 95

```

```

Gln Ser Ala Pro Gly Thr Ile Arg Gly Asp Phe Ala Ile Asp Met Gly
100 105 110

```

```

Arg Asn Val Cys His Gly Ser Asp Ser Val Glu Ser Ala Asn Lys Glu
115 120 125

```

```

Ile Asp Leu Trp Phe Lys Lys Glu Glu Leu Val Glu Tyr Lys Pro Ala
130 135 140

```

```

Leu Phe Gly Trp Ile Tyr Glu
145 150

```

&lt;210&gt; 15

&lt;211&gt; 486

&lt;212&gt; DNA

&lt;213&gt; Candida albicans

&lt;400&gt; 15

```

atgtctcaat tttagcaatt agctccaaaa gacgccaaag gtgaaccata cccatttgaa 60
caattgaaag ggaaagttgt ccttatcgtc aatgttgctt ccaaatgtgg attcactcct 120
caatacaagg gtttagaaga attgaataag aaatttgctg atcaaccagt acaaattctg 180
ggtttcccat gtaatcaatt tggccaccaa gaaccaggta gtaacgaaga aattggatca 240
ttctgttcat tgaactacgg tggttacatt ccagtcttgg ataaaattga agtcaatggg 300
gacaataccg atccagttta taaatatattg aaatcacaaa agagtgggtg tttgggattg 360
accagaatta aatggaattt tgaaaaattc ttgattgacc aaaaatggtg agttattgaa 420
agattcagtt cattgactag tccagaaagt atcggtagca agattgaaga attggtgaag 480
aaataa 486

```

<210> 16  
 <211> 161  
 <212> PRT  
 <213> Candida albicans

<400> 16

Met Ser Gln Phe Tyr Glu Leu Ala Pro Lys Asp Ala Lys Gly Glu Pro  
 1 5 10 15

Tyr Pro Phe Glu Gln Leu Lys Gly Lys Val Val Leu Ile Val Asn Val  
 20 25 30

Ala Ser Lys Cys Gly Phe Thr Pro Gln Tyr Lys Gly Leu Glu Glu Leu  
 35 40 45

Asn Lys Lys Phe Ala Asp Gln Pro Val Gln Ile Leu Gly Phe Pro Cys  
 50 55 60

Asn Gln Phe Gly His Gln Glu Pro Gly Ser Asn Glu Glu Ile Gly Ser  
 65 70 75 80

Phe Cys Ser Leu Asn Tyr Gly Val Thr Phe Pro Val Leu Asp Lys Ile  
 85 90 95

Glu Val Asn Gly Asp Asn Thr Asp Pro Val Tyr Lys Tyr Leu Lys Ser  
 100 105 110

Gln Lys Ser Gly Val Leu Gly Leu Thr Arg Ile Lys Trp Asn Phe Glu  
 115 120 125 130

Lys Phe Leu Ile Asp Gln Asn Gly Lys Val Ile Glu Arg Phe Ser Ser  
 130 135 140

Leu Thr Ser Pro Glu Ser Ile Gly Thr Lys Ile Glu Glu Leu Leu Lys  
 145 150 155 160

Lys

<210> 17  
 <211> 1080  
 <212> DNA  
 <213> Candida albicans

<400> 17

atggctcctc cagcagtttt aagtaaattcc ggtgttatct acggtaaaga cgtcaaagac 60  
 ttgtttgact atgctcaaga aaaaggtttt gccattccag ctatcaatgt cacttcatcc 120  
 tcaactgttg ttgctgcttt agaagctgcc agagacaaca aggtccaat catcttgcaa 180  
 acttctcaag gtggtgctgc ctactttgcc ggtaaaagtg tcgacaacaa agatcaagct 240  
 gcttccattg ctggttcaat tgctgccgct cactacatta gagccattgc tccaacttat 300  
 ggtatcccag ttgttttaca cactgatcac tgtgccaaa aattattgcc atggtttgat 360  
 ggtatgttga aagccgatga agaattcttt gctaagaccg gtactccatt gttctcatcc 420  
 cacatgttggt atttatctga agaaaccgat gacgaaaaca ttgctacttg tgccaaatat 480  
 ttcgaaagaa tggctaaaaa gggtaaatgg ttagaaatgg aaattggtat cactggtggt 540  
 gaagaagatg gtgtcaacaa cgaacacggt gaaaaagatg ctttatacac ttctccagaa 600  
 actgttttcg ctgtctacga atctttacac aagatttctc caaacttttc tattgctgct 660  
 gcttttggtg acgtccacgg tgtttacaac ccaggtaatg tgcaattgag accagaaatc 720  
 ttgggtgacc accaagttta cgctaagaaa caaatttggt ctgatgctaa acaccatta 780

tacttggttt tccacggtgg ttctggttct actcaagaag aattcaacac tgctatcaag 840  
aatggtggtg tcaaggtcaa cttggacact gattgtcaat acgcttactt gactggtatc 900  
agagattacg tcaccaacaa gattgaatac ttgaaagcac cagttggtaa cccagaaggt 960  
gctgacaaac caaacaagaa atactttgac ccaagagtct gggttagaga aggtgaaaag 1020  
accatgtcca agagaattgc tgaagctttg gatattttcc acaccaaagg acaattgtaa 1080

&lt;210&gt; 18

&lt;211&gt; 359

&lt;212&gt; PRT

&lt;213&gt; Candida albicans

&lt;400&gt; 18

Met Ala Pro Pro Ala Val Leu Ser Lys Ser Gly Val Ile Tyr Gly Lys  
1 5 10 15

Asp Val Lys Asp Leu Phe Asp Tyr Ala Gln Glu Lys Gly Phe Ala Ile  
20 25 30

Pro Ala Ile Asn Val Thr Ser Ser Ser Thr Val Val Ala Ala Leu Glu  
35 40 45

Ala Ala Arg Asp Asn Lys Ala Pro Ile Ile Leu Gln Thr Ser Gln Gly  
50 55 60

Gly Ala Ala Tyr Phe Ala Gly Lys Gly Val Asp Asn Lys Asp Gln Ala  
65 70 75 80

Ala Ser Ile Ala Gly Ser Ile Ala Ala Ala His Tyr Ile Arg Ala Ile  
85 90 95

Ala Pro Thr Tyr Gly Ile Pro Val Val Leu His Thr Asp His Cys Ala  
100 105 110

Lys Lys Leu Leu Pro Trp Phe Asp Gly Met Leu Lys Ala Asp Glu Glu  
115 120 125

Phe Phe Ala Lys Thr Gly Thr Pro Leu Phe Ser Ser His Met Leu Asp  
130 135 140

Leu Ser Glu Glu Thr Asp Asp Glu Asn Ile Ala Thr Cys Ala Lys Tyr  
145 150 155 160

Phe Glu Arg Met Ala Lys Met Gly Gln Trp Leu Glu Met Glu Ile Gly  
165 170 175

Ile Thr Gly Gly Glu Glu Asp Gly Val Asn Asn Glu His Val Glu Lys  
180 185 190

Asp Ala Leu Tyr Thr Ser Pro Glu Thr Val Phe Ala Val Tyr Glu Ser  
195 200 205

Leu His Lys Ile Ser Pro Asn Phe Ser Ile Ala Ala Ala Phe Gly Asn  
210 215 220

Val His Gly Val Tyr Lys Pro Gly Asn Val Gln Leu Arg Pro Glu Ile  
225 230 235 240

Leu Gly Asp His Gln Val Tyr Ala Lys Lys Gln Ile Gly Thr Asp Ala  
245 250 255

Lys His Pro Leu Tyr Leu Val Phe His Gly Gly Ser Gly Ser Thr Gln  
260 265 270



Glu Glu Phe Asn Thr Ala Ile Lys Asn Gly Val Val Lys Val Asn Leu  
275 280 285

Asp Thr Asp Cys Gln Tyr Ala Tyr Leu Thr Gly Ile Arg Asp Tyr Val  
290 295 300

Thr Asn Lys Ile Glu Tyr Leu Lys Ala Pro Val Gly Asn Pro Glu Gly  
305 310 315 320

Ala Asp Lys Pro Asn Lys Lys Tyr Phe Asp Pro Arg Val Trp Val Arg  
325 330 335

Glu Gly Glu Lys Thr Met Ser Lys Arg Ile Ala Glu Ala Leu Asp Ile  
340 345 350

Phe His Thr Lys Gly Gln Leu  
355